

АНАЛИЗ ИЗМЕНЕНИЙ ТОЧНОСТИ ПРОГНОЗА ГЕНЕТИЧЕСКОЙ ЦЕННОСТИ ПО РАЗЛИЧНЫМ СТАТИСТИЧЕСКИМ МОДЕЛЯМ И ПРИ РАЗЛИЧНОМ СООТНОШЕНИИ ВАРИАНС

Н. М. ХРАМЧЕНКО, А. В. РОМАНЕНКО

*РУП «Научно-практический центр Национальной академии наук по животноводству»,
г. Жодино, Республика Беларусь, 222160*

(Поступила в печать 30.01.2019 г.)

В статье приведен анализ изменений точности генетической (племенной) оценки селекционируемых признаков свиней в зависимости от используемой статистической модели и величины коэффициента наследуемости. На основании исследований установлено, что точность прогноза в большей степени зависит не от использованной для расчета племенной ценности модели, а от количества информации и пропорциональна коэффициенту наследуемости.

Ключевые слова: *свиньи, генетическая оценка, статистическая модель, наследуемость, точность (надежность) расчета племенной ценности, эффект селекции.*

The paper presents analysis of variability in the accuracy of genetic (breeding) assessment of breeding traits of pigs depending on the statistical model applied and heritability coefficient value. It has been determined based on the research that prediction accuracy depends to a greater extent not on the model used for breeding value calculation, but on the data volume, and it is proportional to the heritability coefficient.

Key words: *pigs, genetic assessment, statistical model, heritability, accuracy (reliability) of breeding value calculation, breeding progress.*

Введение. Современные селекционные программы стран с развитым свиноводством, ставящие своей целью генетическое улучшение экономически значимых признаков, основаны на законах популяционной генетики и генетики количественных признаков. Поэтому четкое понимание генетических принципов, лежащих в основе таких программ, является основой принятия грамотных селекционных решений. При использовании методологии BLUP точность расчета прогнозируемой племенной ценности (Reliability, REL) является критерием использования животного, учитывающим риски ухудшить племенные качества стада, а не критерия выбора того или иного производителя.

Анализ источников. На точность оценки генетической ценности влияют следующие факторы:

наследуемость: чем выше наследуемость, тем выше точность оценки; однако при этом преимущества более точных методов оценки (селекционный индекс, BLUP) перед оценкой по фенотипу снижаются;

общий объем данных (количество животных и наблюдений, а также число поколений);

достоверность и точность индивидуального учета продуктивности животных;

ошибки родословных;

генетические связи между животными (наблюдениями), относящимися к разным уровням фиксированных факторов в модели, в первую очередь между так называемыми единицами менеджмента;

объем данных для конкретного животного (число наблюдений и/или признаков, число сибсов, полусибсов, потомков и т.д.);

корректность и полнота модели: в модель следует включить все фиксированные и случайные факторы, которые оказывают достоверное влияние на анализируемый признак;

точность селекционно-генетических параметров (наследуемость, повторяемость, генетические и средовые корреляции): если используемые значения отличаются от истинных, то фактические дисперсии ошибок прогноза генетической ценности будут больше, чем рассчитанные [1].

Мерой эффективной оценки племенной ценности является достоверность (Reliability, REL), которая является мерой количества, а не качества информации. На сегодня нет такого метода, который мог бы выявить подтасовку или искажение первичных данных. Компьютеры могут оказать большую помощь при принятии правильных решений по выбору животных с лучшими генотипами, но необходимо быть уверенными в корректности исходных данных [2].

Цель работы – проведение анализа изменений прогноза генетической ценности и его точности при различных статистических моделях и соотношении вариантов (наследуемости) по селектируемым признакам.

Материал и методика проведения исследований. Исследования проводились на основе метода BLUPAM на базе данных опытной популяции свиней, сгенерированной из базы данных государственной информационной системы в области племенного свиноводства.

Точность (r) прогноза – это корреляция между истинными и прогнозируемыми племенными оценками [1]. При оценке методом BLUP точность обычно выражается как надежность: квадратичная корреляция между истинными и предсказанными племенными оценками (r^2). Расчеты для r или r^2 требуют диагональных элементов инвертированной ММЕ, как показано Хендерсоном (1975) [3].

При использовании BLUP AM уравнение смешанной модели ММЕ имеет вид [4]:

$$\begin{bmatrix} \mathbf{X}'\mathbf{X} & \mathbf{X}'\mathbf{Z} \\ \mathbf{Z}'\mathbf{X} & \mathbf{Z}'\mathbf{Z} + \mathbf{A}^{-1}\alpha \end{bmatrix} \begin{bmatrix} \hat{\mathbf{b}} \\ \hat{\mathbf{a}} \end{bmatrix} = \begin{bmatrix} \mathbf{X}'\mathbf{y} \\ \mathbf{Z}'\mathbf{y} \end{bmatrix}$$

матрица коэффициентов представлена как: $\begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix}$

и обобщенная обратная матрица коэффициентов: $\begin{bmatrix} C^{11} & C^{12} \\ C^{21} & C^{22} \end{bmatrix}$.

Варианса ошибки прогноза (Prediction Error Variance, PEV):

$$PEV = \text{var}(a-\hat{a}) = C^{22} \sigma_e^2$$

Для расчета PEV для животных необходимы диагональные элементы матрицы коэффициентов уравнений животных. PEV можно рассматривать как долю аддитивной генетической вариации, не учитываемую прогнозом. Поэтому:

$$PEV = C^{22} \sigma_e^2 = (1-r^2) \sigma_a^2,$$

где r^2 – квадратная корреляция между истинными и рассчитанными племенными оценками.

Таким образом, ошибка прогноза имеет нормальное распределение со средней равной 0 и дисперсией $(1-r^2) \sigma_a^2$:

ошибка прогноза $\sim N(0, (1-r^2) \sigma_a^2)$.

Корень квадратный из PEV дает стандартную ошибку прогноза (StandardErrorPrediction, SEP):

$$SEP = \sqrt{PEV} = \sigma_a \sqrt{1-r^2}.$$

Из этого следует, что для снижения ошибки прогноза необходимо использовать такие доступные методы, которые максимизируют r при имеющемся количестве информации.

Надежность (Reliability, REL) оценки для животного рассчитывается как:

$$REL = r^2 = 1 - \frac{PEV_i}{\sigma_a^2}.$$

Для сравнения точности прогноза по различным статистическим моделям использовались следующие биометрические модели смешанного типа:

для среднесуточного прироста содержания постного мяса и толщины шпика:

$y = B+S+F+a+e$ (оптимальная):

$$y = S+F+a+e$$

$$y = B+F+a+e$$

$$y = B+S+a+e$$

$$y = B+a+e$$

$$y = S+a+e$$

$$y = F+a+e,$$

для многоплодия свиноматок и массы гнезда к отъему:

$$y = B+F+NF+a+e \text{ (оптимальная):}$$

$$y = F+NF+a+e$$

$$y = B+NF+a+e$$

$$y = NF+a+e,$$

где: y – вектор наблюдаемой зависимой переменной, B – фиксированный фактор «порода», S – фиксированный фактор «пол», F – фиксированный фактор «хозяйство», NF – фиксированный фактор «номер».

ра опороса», а – случайный фактор «животное», е – вектор случайных остаточных (неучтенных) эффектов.

Коэффициент наследуемости (h^2) [5] рассчитывался по формуле:

$$h^2 = \frac{\sigma_a^2}{\sigma_a^2 + \sigma_e^2},$$

где: σ_a^2 – изменчивость, обусловленная генотипом животного (аддитивная вариация); σ_e^2 – остаточная изменчивость (случайная вариация).

Вариансы σ_a^2 и σ_e^2 получены из результатов оценки племенной ценности селекционируемых признаков опытной популяции животных рассчитанной по оптимальным моделям.

Для расчета вариационно-ковариационных взаимодействий рандомизированных факторов модели и решения смешанных линейных уравнений на основе BLUP использовался пакет программ BLUPF90 [6].

Для расчета генетического превосходства потомства использовали следующее уравнение [7]:

$$\hat{S} = r_{gl} \frac{\sigma_g}{\sigma_l} (i\sigma_l) = \frac{ir_{gl}\sigma_g}{L}.$$

В наших исследованиях применялся частный случай усеченной селекции. При этой форме отбора для разведения отбирались все животные, больше определенного значения индексах (рис. 1). Точка усечения определялась p -долей отбираемых для селекции животных. Предполагалось, что значения прогноза племенной ценности нормально распределены, отношение между p , x (измеренными в единицах стандартного отклонения) и i получено из свойств нормального распределения, L – генерационный интервал (в нашем случае принят равным 1 году):

$$i = z/p,$$

где: z – высота нормального распределения в точке усечения x и задавался формулой: $z = \frac{e^{-1/2x^2}}{\sqrt{2\pi}}$ где π , до 9 знаков – 3,141592654.

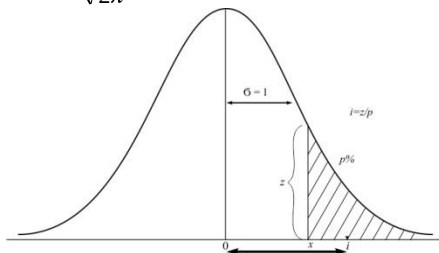


Рис. 1. Расчет интенсивности селекции из стандартного нормального распределения

Результаты исследований и их обсуждение. При анализе зависимости точности рассчитанных значений племенной ценности от используемой статистической модели установлено, что включение фиксированных факторов в модели признаков собственной продуктивности не оказало значительного влияния на величины аддитивной и случайной дисперсий (табл. 1). Наследуемость исследуемых признаков находилась в пределах: по среднесуточному приросту – 0,55–0,62, по содержанию мяса в туше – 0,78–0,85, по толщине шпика – 0,71–0,77; что соответствует теории наследования данных признаков. Необходимо отметить, что фактор «хозяйство» оказывал значительное влияние на величину аддитивной дисперсии. Исключение его из моделей повысило коэффициент наследуемости на 0,04–0,07. Это свидетельствует о том, что доля изменчивости, обусловленная фактором «хозяйство», включалась в аддитивную дисперсию.

Аналогичная тенденция наблюдалась по массе гнезда к отъему: величина наследуемости после исключения фактора «хозяйство» увеличилась на 0,07 единицы. На величину наследуемости многоплодия влияние факторов модели не выявлено: она находилась на уровне 0,14–0,15.

Таблица 1. Анализ соотношения аддитивной и случайной дисперсий исследуемых моделей

модель		ССП	СПМ	ТШ	модель		МГ	М
y = B+S+F+a+ e	σ_a^2	4353	4,384	3,061	y = B+F+NF+ a+e	σ_a^2	54,65	1,139
	σ_e^2	3549	1,257	1,266		σ_e^2	436,7	6,994
	h^2	0,56	0,78	0,71		h^2	0,11	0,14
y = F+NF+a+e	σ_a^2	4372	4,398	3,073	y = F+NF+a+e	σ_a^2	61,46	1,205
	σ_e^2	3557	1,260	1,280		σ_e^2	434,6	6,980
	h^2	0,55	0,78	0,70		h^2	0,12	0,15
y = B+F+a+e	σ_a^2	4414	4,430	3,064	y = B+NF+a+e	σ_a^2	98,74	1,153
	σ_e^2	3542	1,242	1,268		σ_e^2	424,8	6,990
	h^2	0,56	0,78	0,71		h^2	0,19	0,14
y = B+S+a+e	σ_a^2	5219	5,215	3,626	y = NF+a+e	σ_a^2	98,59	1,248
	σ_e^2	3247	1,025	1,080		σ_e^2	425,0	6,969
	h^2	0,62	0,84	0,77		h^2	0,19	0,15
y = B+a+e	σ_a^2	5201	5,38	3,651				
	σ_e^2	3295	0,98	1,072				
	h^2	0,61	0,85	0,77				
y = S+a+e	σ_a^2	5245	5,248	3,657				
	σ_e^2	3256	1,028	1,080				
	h^2	0,62	0,84	0,77				
y = F+a+e	σ_a^2	4423	4,44	3,082				
	σ_e^2	3551	1,25	1,280				
	h^2	0,55	0,78	0,71				

Здесь и далее: y – вектор наблюдаемой зависимой переменной, B – фиксированный фактор «порода», S – фиксированный фактор «пол», F – фиксированный фактор «хозяйство», NF – фиксированный фактор «номера опороса», a – случайный фактор «животное», e – вектор случайных остаточных (неучтенных) эффектов.

Точность рассчитанных значений племенной ценности (табл. 2) по признакам собственной продуктивности находилась на уровне 0,63–0,69 – по среднесуточному привесу; 0,80–0,85 – по содержанию мяса в туше и 0,74–0,80 – по толщине шпика, незначительно увеличиваясь у моделей с отсутствием фиксированного фактора «хозяйство». Это связано с тем, что точность прогноза в большей степени зависит не от модели, а от количества информации о родственниках животного в опытной популяции.

Таблица 2. Точность прогноза рассчитанной племенной ценности исследуемых моделей

Модель	ССП	СПМ	ТШ	Модель	МГ	М
Точность оценки BLUP						
$y = B + S + F + a + e$	0,63	0,80	0,74	$y = B + F + NF + a + e$	0,40	0,44
$y = F + NF + a + e$	0,63	0,80	0,74	$y = F + NF + a + e$	0,42	0,46
$y = B + F + a + e$	0,63	0,80	0,75	$y = B + NF + a + e$	0,51	0,45
$y = B + S + a + e$	0,68	0,85	0,80	$y = NF + a + e$		0,47
$y = B + a + e$	0,68	0,85	0,80			
$y = S + a + e$	0,69	0,85	0,80			
$y = F + a + e$	0,64	0,80	0,75			

Анализ корреляции показал, что наибольшая взаимосвязь с фактической продуктивностью получена при использовании для расчета племенной ценности простых моделей. Это вызвано тем, что на расчет племенной ценности фиксированные факторы не оказывали влияния, в результате чего расчетная оценка ближе к фактической (фенотипу).

При проведении исследований по оценке влияний наследуемости на точность генетической оценки, для расчета вариантов были использованы биометрические оптимальные модели смешанного типа: для среднесуточного прироста, толщины шпика и содержания постного мяса $y = B + S + a + e$; для многоплодия свиноматок и массы гнезда к отъему $y = B + F + NF + a + e$, где y – вектор наблюдаемой зависимой переменной, B – фиксированный фактор «порода», S – фиксированный фактор «пол», F – фиксированный фактор «хозяйство», NF – фиксированный фактор «номер опороса», a – случайный фактор «животное», e – вектор случайных остаточных (неучтенных) эффектов. Расчет вариантов и прогноз племенной ценности по селекционируемым признакам проводился на основе метода BLUPAM.

Для расчета точности (надежности) получены следующие аддитивные варианты: среднесуточный прирост (ССП) – 4353 (h^2 -0.551), содержание постного мяса (СПМ) – 4,381 (h^2 -0.777), толщина шпика (ТШ) – 3,061 (h^2 -0.707), многоплодие (М) – 1,139 (h^2 -0.140), масса гнезда при отъеме (МГ) – 58,07 (h^2 -0.124), в скобках указаны фактические показатели наследуемости, полученные при анализе опытной популяции животных. Изменение наследуемости проводилось путем изменения случайной вариации. Таким образом, только соотношение

варианс (наследуемость) оказывало влияние на изменение надежности прогноза селекционируемых признаков.

Анализ зависимости прогноза генетической ценности от наследуемости показал, что надежность прогноза генетической ценности, рассчитанной на основе показателей измерений признаков и родственных связей животных опытной популяции свиней по признакам собственной продуктивности, возрастает по мере увеличения коэффициентов наследуемости: по признакам собственной продуктивности – от 0,21 до 0,99, по репродуктивным признакам – от 0,38 до 0,98 (табл. 3).

Таблица 3. Зависимость точности (надежности) прогноза генетической ценности при различном соотношении варiances (наследуемости) селекционируемых признаков

Признак \ h^2	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9	1,0
Признак	0,23	0,35	0,44	0,52	0,59	0,67	0,74	0,82	0,90	0,99
СПМ	0,21	0,33	0,42	0,50	0,58	0,66	0,74	0,81	0,90	0,99
ТШ	0,23	0,35	0,44	0,52	0,59	0,67	0,74	0,82	0,90	0,99
М	0,38	0,51	0,60	0,67	0,73	0,78	0,83	0,88	0,93	0,98
МГ	0,38	0,51	0,60	0,66	0,73	0,78	0,83	0,88	0,93	0,98

Установлено, что при низких коэффициентах наследуемости на точность прогноза первоочередное влияние оказывает количество привлеченных данных о предках. Так, точность прогноза репродуктивных признаков при коэффициенте наследуемости 0,1 была выше, чем при прогнозе признаков оценки собственной продуктивности в среднем на 0,16, что объясняется большим количеством данных об опоросах пробандов и их родственниках, используемых при прогнозе. При увеличении коэффициентов наследуемости точность прогноза стремится к 100 %.

Взаимосвязь наследуемости и точности прогноза (рис. 2) имеет зависимость, близкую к линейной. Угол наклона графиков функций определяется изменениями в точности прогноза за счет привлечения данных предков при низких коэффициентах наследуемости признаков в популяции.

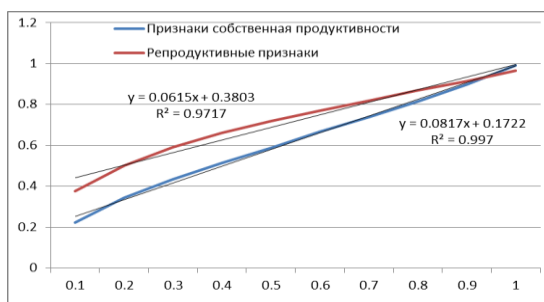


Рис. 2. Графики взаимосвязи наследуемости и точности прогноза по исследуемым признакам

В свиноводстве цель селекционно-племенной работы заключается в генетическом улучшении популяции. Генетическое улучшение достигается путем выбора лучших особей текущего поколения и использования их в качестве родителей следующего.

Цель генетического улучшения поголовья свиней состоит в том, чтобы увеличить генетический потенциал животных по селекционируемым признакам, путем генетического отбора для достижения общей экономической цели – увеличить прибыль на животное в год.

Разработка и внедрение в практику методик прогнозирования генетического превосходства родителей для получения животных новых генераций является частью программ разведения состоящих их из следующих компонентов[7]:

- 1) система регистрации данных о племенных животных (кандидаты на отбор);
- 2) методики оценки (прогнозирования) племенной генетической ценности животных;
- 3) система подбора и отбора животных следующего поколения;
- 4) структура для распространения генетического улучшения программы разведения в производственной популяции.

Точность (надежность) является одним из основных факторов, влияющих на темпы улучшения популяций, заложенные в программах селекции. Для прогноза генетического тренда в наших исследованиях мы использовали три уровня отбора: 10, 30 и 50 % лучших животных по каждому исследуемому признаку. Установлено (табл. 4), что в анализируемой популяции отбор 10 % лучших животных позволит получить заметное увеличение исследуемых признаков у потомков: среднесуточный прирост – на 91,2 грамм, многоплодие – на 1,31 поросенка.

Таблица 4. Прогнозируемый генетический тренд следующего поколения

p%	i	x	Среднесуточный прирост		Многоплодие	
			r^2	Генетическое превосходство родителей (S)	r^2	Генетическое превосходство родителей (S)
10	1,755	1,282	0,62	91,2	0,49	1,31
30	1,159	0,524	0,62	60,2	0,47	0,85
50	0,798	0,000	0,62	41,5	0,46	0,58

При отборе 30 и 50 % лучших животных темпы селекции по исследуемым признакам ожидаемо снижаются, однако остаются высокими, что свидетельствует о высокой вариабельности оцениваемых признаков, позволяющей поддерживать высокие темпы селекции.

Необходимо уточнить, что данный тренд рассчитан по каждому признаку в отдельности. В настоящее время такой вид селекции практически не используется, так как в основе программ разведения лежат мультипризнаковые селекционные индексы. Весовой вклад каждого

селекционируемого признака определяется экономическим весом, генетической и фенотипической взаимосвязью, наследуемостью и изменчивостью в популяции.

Заключение. Установлено, что использование коэффициентов наследуемости и корреляции с фактической продуктивностью при оценке прогноза может привести к ложным ассоциациям и исключить наиболее влияющие на продуктивность факторы, в нашем случае фиксированный фактор «хозяйство». Точность прогноза в большей степени зависит не от использованной для расчета племенной ценности модели, а от количества информации. В нашем случае она находилась на уровне 0,63–0,69 – по среднесуточному привесу; 0,80–0,85 – по содержанию мяса в туше и 0,74–0,80 – по толщине шпика; 0,40–0,51 – по массе гнезда при отъеме; 0,44–0,47 – по многоплодию, незначительно увеличиваясь у моделей с отсутствием фиксированного фактора «хозяйство».

Анализ изменений прогноза генетической ценности и его точности при различном соотношении вариантов (наследуемости) селекционируемых признаков показал, что с увеличением наследуемости надежность прогноза и взаимосвязь с фактической продуктивностью стремится к 100 %, что обусловлено увеличением влияния собственной продуктивности на генотип оцениваемых животных. Следовательно, использование метода BLUP наиболее эффективно для прогноза племенной ценности по признакам, имеющим низкую наследуемость. С увеличением наследуемости эффективность данного метода снижается. Установлено, что в исследуемой популяции отбор 10 % лучших животных для получения следующего поколения позволит получить статистически значимое увеличение всех исследуемых признаков продуктивности: среднесуточного прироста на 108 грамм, содержания постного мяса на 0,12 %, толщины шпика на 0,08 мм, массы гнезда на 0,99 кг, многоплодия на 0,01 поросенка.

ЛИТЕРАТУРА

1. Даншин, В. А. Оценка генетической ценности животных / В. А. Даншин. – Киев : Аграрна наука, 2008. – 180 с.
2. Кузнецов, В. М. Методы племенной оценки животных с введением в теорию BLUP / В. М. Кузнецов. – Киров : Зональный НИИСК Северо-Востока, 2003. – 358 с.
3. Henderson, C. R. Best linear unbiased estimation and prediction under a selection model / C. R. Henderson // *Biometrics*. – 1975. – Vol. 31. – P. 423–447.
4. Mrode, R. A. Linear models for the prediction of animal breeding values / R. A. Mrode. – 2nd ed. – Wallingford, 2005. – 368 p.
5. Крюков, В. И. Генетика. Часть 14. Генетика количественных признаков и генетические основы селекции : учеб. пособие для вузов / В. И. Крюков. – Изд. 2-е., доп. и испр. – Орёл : изд-во ОрёлГАУ, 2011. – 134 с.
6. Manual for BLUPF90 family of programs / I. Misztal [et. al.]; University of Georgia. – Athens, USA, 2015.
7. Dekkers, J. C. M., Gibson, J. P., Bijma, P., and van Arendonk, J. A. M. 2005. Design and optimization of animal breeding programmes. Iowa State University. <http://www.anslab.iastate.edu/Class/AnS652X/>.