

ارزیابی دام ها با استفاده از داده های مختلف و صحت ارزیابی ها

شاخص انتخاب

تئوری شاخص انتخاب برای اولین بار در دهه‌های ۱۹۳۰ و ۱۹۴۰ به عنوان یک روش پیش بینی ژنتیکی و ابزاری برای صفات ترکیبی برای انتخاب حیوانات به روش مطلوب اقتصادی، مطرح شد. شاخص انتخاب در اصل به عنوان ترکیب خطی یا شاخصی از انواع مختلف اطلاعات فنوتیپی و عوامل تصحیح مناسب بوده و به صورت زیر است:

$$I = b_1x_1 + b_2x_2 + \dots + b_nx_n$$

که I = مقدار شاخص یا پیش بینی ژنتیکی

b_i = عامل تصحیح

x_i = بخشی از اطلاعات فنوتیپی (رکورد عملکرد یا میانگین گروهی از رکوردهای عملکرد)

n = تعداد کل اطلاعات فنوتیپی

در شاخص انتخاب (یا در حالت کلی، در همه روش‌های پیش بینی ژنتیکی)، اطلاعات مورد استفاده برای محاسبه پیش بینی ژنتیکی حیوانات از سه منبع به دست می‌آیند:

(۱) رکورد (های) عملکرد خود فرد

(۲) رکوردهای حاصل از اجداد و (یا) خویشاوندان جانبی فرد (اطلاعات شجره‌ای)

(۳) رکوردهای عملکرد حاصل از نتاج فرد (اطلاعات نتاج).

مقادیر نسبی داده‌های حاصل از این منابع متفاوت است. حیوانات متولد نشده یا بسیار جوان تنها اطلاعات شجره‌ای دارند و با افزایش سن آنها، اطلاعات رکورد خود حیوان به دست می‌آید. اگر حیوانات به عنوان والدین انتخاب شوند، تعداد زیادی فرزند تولید می‌کنند که حجم زیادی از اطلاعات نتاج را خواهند داشت. رکوردهای حاصل از هر سه منبع می‌تواند رکوردهای یک صفت خاص باشد که برای پیش بینی مورد محاسبه قرار می‌گیرد، یا می‌تواند شامل رکوردهای دیگری (صفات همبسته از نظر ژنتیکی) باشند. به عنوان مثال، در گونه‌ای که سخت زایی مشکل ساز می‌باشد، پیش بینی استعداد ژنتیکی به سخت زایی را می‌توان از طریق اندازه‌گیری مستقیم سخت زایی مانند امتیاز بندی سخت زایی و یا از طریق وزن تولد (یک صفت همبسته از نظر ژنتیکی) و یا هر دو صفت تعیین نمود. اطلاعات مورد استفاده در شاخص انتخاب $-Xi$ ها- از طریق منابع زیادی نظیر

اطلاعات رکورد خود حیوان، شجره و اطلاعات نتاج به دست می‌آید. به عنوان مثال، ممکن است x_1 عملکرد خود حیوان برای یک صفت، x_2 میانگین عملکرد برادر و خواهران ناتنی حیوان برای همان صفت، x_3 میانگین عملکرد نتاج فرد و x_4 ، x_5 و x_6 به ترتیب نشان دهنده عملکرد خود فرد، برادر و خواهران ناتنی و نتاج فرد برای یک صفت همبسته باشد. در هر حالت x یک عدد منفرد است. در شاخص انتخاب، هر بخش از اطلاعات فنوتیپی به طور معمول به صورت انحراف از میانگین جامعه و یا گروه هم عصر بیان می‌شوند. شاخص‌های انتخاب باید تنها زمانی برای پیش بینی ژنتیکی استفاده شوند که اطلاعات عملکرد از گروه‌های هم عصر که از لحاظ ژنتیکی مشابه هستند، به دست آمده باشد. در عمل، به این معنی است که شاخص انتخاب احتمالاً باید در داخل افراد گله به کار برده شود و نباید بین جمعیت‌ها استفاده شود و اگر گله خاصی در طول زمان پیشرفت ژنتیکی قابل ملاحظه ای داشته باشد، بنابراین اطلاعات گروه‌های هم عصر مسن تر (گروه‌های مسن تر و بنابراین میانگین ارزش‌های اصلاحی متفاوت) باید از شاخص حذف شوند.

شاخص‌های انتخاب ممکن است از نظر کاربرد محدودیت داشته باشند اما با وجود این، بسیار مفید هستند. به عنوان مثال، فرض کنید یک پرورش دهنده گوسفند می‌خواهد قوچ‌های خود را بر اساس عملکرد نتاج در داخل گله خود مقایسه کند. با استفاده از شاخص انتخاب و بر اساس اطلاعات نتاج می‌توان EBV یا EPD هر قوچ را محاسبه کرد.

ساده‌ترین شاخص انتخاب، حالتی است که تنها از یک منبع اطلاعات (یک x) استفاده شود. این شاخص به صورت زیر است:

$$I = bx$$

از آنجا که تنها یک معادله برای محاسبه ضریب تابعیت b نیاز است، محاسبه این معادله آسان خواهد بود.

فرمول‌های ضرایب تابعیت مورد استفاده برای محاسبه برخی از شاخص‌های انتخاب حاوی یک منبع اطلاعات، در جدول ۱ ارائه شده است. اولین ستون سمت راست جدول، شامل نوع ارزش پیش بینی شده -EBV، EPD، یا MPPA است و ستون بعدی نوع ارزش حقیقی که مورد پیش بینی قرار می‌گیرد را نشان می‌دهد. ستون سوم، یک منبع اطلاعات (x) یا معیار مورد استفاده را توصیف می‌کند، و ستون بعدی فرمول‌های ضرایب تابعیت متناسب (b) را نشان می‌دهد. ستون آخر فرمول‌های صحت پیش بینی -همبستگی بین ارزش حقیقی (BV، PD یا PA) و پیش بینی شده آن‌ها (ارزش شاخص I) را نشان می‌دهد.

جدول ۱- نمونه‌ای از فرمول‌های ضرایب تابعیت مورد استفاده برای محاسبه پیش بینی‌های حاصل از یک منبع اطلاعات. و صحت‌های همبسته

پیش بینی (I)	ارزش حقیقی	منبع اطلاعات (X)	ضریب تابعیت (b)	صحت
EBV	BV	P: یک رکورد (غیر تکراری) عملکرد از حیوان	h^2	h
EBV	BV	\bar{P} : میانگین n رکورد از حیوان	$\frac{nh^2}{1 + (n - 1)r}$	$\sqrt{\frac{nh^2}{1 + (n - 1)r}}$
MPPA	PA	\bar{P} : میانگین n رکورد از حیوان	$\frac{nr}{1 + (n - 1)r}$	$\sqrt{\frac{nr}{1 + (n - 1)r}}$
EBV	BV	\bar{P} : میانگین یک رکورد از m ناتنی	$\frac{mh^2}{4 + (m - 1)h^2}$	$\sqrt{\frac{\frac{1}{4}mh^2}{4 + (m - 1)h^2}}$
EPD	PD	\bar{P} : میانگین یک رکورد از m ناتنی	$\frac{\frac{1}{2}mh^2}{4 + (m - 1)h^2}$	$\sqrt{\frac{\frac{1}{4}mh^2}{4 + (m - 1)h^2}}$
EBV	BV	\bar{P} : میانگین یک رکورد از m تنی	$\frac{mh^2}{2 + (m - 1)(h^2 + 2c_{FS}^2)}$	$\sqrt{\frac{\frac{1}{2}mh^2}{2 + (m - 1)(h^2 + 2c_{FS}^2)}}$
EBV	BV	\bar{P} : میانگین رکوردهای تنها از p نتاج	$\frac{2ph^2}{4 + (p - 1)h^2}$	$\sqrt{\frac{ph^2}{4 + (p - 1)h^2}}$
EPD	PD	\bar{P} : میانگین یک رکورد از p نتاج	$\frac{ph^2}{4 + (p - 1)h^2}$	$\sqrt{\frac{ph^2}{4 + (p - 1)h^2}}$
EBV	BV	\bar{P} : میانگین یک رکورد از هر نتاج از 1 تعداد توله در k زایش	$\frac{2lkh^2}{4 + (k - 1)(2h^2 + 4c_{FS}^2) + l}$	$\sqrt{\frac{2lkh^2}{4 + (k - 1)(2h^2 + 4c_{FS}^2) + l}}$
EBV	BV	\bar{P} : میانگین n رکورد از هر حیوان از p نتاج	$\frac{\frac{1}{2}ph^2}{1 + (n - 1)r + (p - 1)\frac{h^2}{4}}$	$\sqrt{\frac{\frac{1}{4}ph^2}{1 + (n - 1)r + (p - 1)\frac{h^2}{4}}}$

برای درک کاربرد جدول ۱، فرض کنید می‌خواهیم MPPA تولید شیر یک گاو شیری را محاسبه کنیم. منبع اطلاعات در این مورد، پیشینه تولید شیر این گاو است. او به عنوان یک گاو مسن دارای پنج دوره شیردهی با میانگین ۱۰۷۲ پوند بالاتر از میانگین‌های گروه هم عصر است. فرمول ضریب تابعیت متناسب آن در ردیف سوم جدول ۱ وجود دارد.

$$\frac{nr}{1 + (n - 1)r}$$

تکرار پذیری (r) تولید شیر به طور معمول در حدود ۰/۵، و تعداد رکوردهای گاو شیری (n) نیز برابر ۵ است. پس:

$$\begin{aligned} b &= \frac{nr}{1 + (n - 1)r} \\ &= \frac{5(0/5)}{1 + (5 - 1)(0/5)} \\ &= 0/833 \end{aligned}$$

به ازای یک پوند افزایش در میانگین ۵ رکورد این گاو شیری انتظار داریم که قابلیت تولید حقیقی این گاو ۰/۸۳۳ پوند افزایش یابد. اکنون برای محاسبه MPPA این گاو، باید این ضریب تابعیت را در شاخص انتخاب جای‌گذاری کنیم. بنابراین:

$$\begin{aligned} I &= bx \\ &= 0/833(+1072) \\ &= +893 \text{ پوند} \end{aligned}$$

MPPA این گاو شیری برابر ۸۹۳+ پوند خواهد بود. به عبارت دیگر، انتظار می‌رود او نسبت به میانگین گاوهای هم عصر خود ۸۹۳ پوند شیر بیشتر تولید کند. اگر به طور معمول میانگین تولید شیر این گله حدود ۱۴۰۰۰ پوند باشد، بنابراین پیش‌بینی می‌شود که رکورد بعدی این گاو شیری ۱۴۸۹۳ پوند شیر باشد. صحت این پیش‌بینی ($r_{PA, \widehat{PA}}$) را می‌توان با استفاده از فرمول ستون آخر جدول ۱ محاسبه نمود. بنابراین:

$$\begin{aligned}
 r_{PA, \widehat{PA}} &= \sqrt{\frac{nr}{1 + (n-1)r}} \\
 &= \sqrt{\frac{5(0/5)}{1 + (5-1)0/5}} \\
 &= \sqrt{0/833} \\
 &= 0/93
 \end{aligned}$$

Regression for amount of information

تابعیت تخمین از میزان اطلاعات

توجه داشته باشید با وجود این که تولید شیر گاو آیریس از میانگین هم عصرهای خود ۱۰۷۲ پوند بیشتر بود، پیش بینی قابلیت تولید او کمتر از ۱۰۷۲ پوند است. این امر به دلیل آن است که پیش بینی گاهی اوقات دستخوش پدیده‌ای به نام تابعیت تخمین از میزان اطلاعات می‌شود. این پدیده یک فرآیند محاسباتی است که باعث می‌شود پیش بینی‌های ژنتیکی کم و بیش محافظه کارانه (نزدیک‌تر به میانگین) شوند و به مقدار اطلاعات مورد استفاده در محاسبات این پیش بینی‌ها بستگی دارد. برای درک بهتر نحوه عمل تابعیت تخمین از میزان اطلاعات، دو گاو دیگر از همان گله قبلی (گاو شماره ۱ و شماره ۲) را در نظر بگیرید. گاو شماره یک دارای دو رکورد شیردهی با میانگین ۱۲۰۴ پوند بالاتر از میانگین گروه هم عصر خود می‌باشد. در مورد او خواهیم داشت:

$$\begin{aligned}
 b &= \frac{nr}{1 + (n-1)r} \\
 &= \frac{2(0/5)}{1 + (2-1)(0/5)} \\
 &= 0/667
 \end{aligned}$$

مقدار MPPA گاو شماره یک برابر است با:

$$\begin{aligned}
 I &= bx \\
 &= 0/667(+1204) \\
 &= +803 \text{ پوند}
 \end{aligned}$$

صحت این پیش بینی برابر است با:

$$\begin{aligned} r_{PA, \hat{PA}} &= \sqrt{\frac{nr}{1 + (n-1)r}} \\ &= \sqrt{0.667} \\ &= 0.82 \end{aligned}$$

گاو شماره دو تنها یک رکورد تولید شیر دارد و مقدار آن ۹۱۸ پوند کمتر از میانگین هم عصرهایش است. در مورد مورد گاو شماره دو داریم:

$$\begin{aligned} b &= \frac{nr}{1 + (n-1)r} \\ &= \frac{1(0.5)}{1 + (1-1)(0.5)} \\ &= 0.5 \end{aligned}$$

MPPA برای گاو شماره دو برابر خواهد بود با:

$$\begin{aligned} I &= bx \\ &= 0.5(-918) \\ &= -459 \text{ پوند} \end{aligned}$$

و صحت پیش بینی برابر:

$$\begin{aligned} r_{PA, \hat{PA}} &= \sqrt{\frac{nr}{1 + (n-1)r}} \\ &= \sqrt{0.5} \\ &= 0.71 \end{aligned}$$

تابعیت پیش بینی‌های ژنتیکی از میزان اطلاعات، منجر به پیش بینی نااریب (Unbiased) خواهد شد. به این معنی است که اطلاعات بیشتر مورد استفاده در پیش بینی‌های بعدی همان دام باعث تغییر پیش بینی‌ها چه در جهت مثبت و یا منفی می‌شود. به علت این که پیش بینی‌های ژنتیکی تابع میزان اطلاعات هستند، در حقیقت برای صحت تصحیح می‌شوند. به عبارت دیگر، صحت پیش بینی برای پیش بینی ارزش مورد نظر نیز در نظر گرفته می‌شود. از مزایای تصحیح برای صحت، فراهم نمودن مقایسه مستقیم پیش بینی‌های حیوانات مختلف بدون در نظر گرفتن صحت این پیش بینی‌ها خواهد بود. به عنوان مثال، اگر دو حیوان دارای EBV مشابه باشند، اما صحت EBV برای یکی از آنها بسیار بالاتر از دیگری باشد، باز هم این دو حیوان را از لحاظ ژنتیکی یکسان در نظر می‌گیریم. البته ارزش اصلاحی حقیقی حیوان با صحت پایین‌تر، به طور معنی داری کمتر از مقدار مورد انتظار است، و انتخاب این حیوان خطر پذیری بیشتری دارد. به عبارت دیگر، احتمال اینکه ارزش اصلاحی حقیقی این حیوانات از نظر معنی داری بهتر از مورد انتظار باشد، بالاتر است. از این رو این حیوان شانس بالاتری برای انتخاب دارد. مزیت ارزش‌های با صحت بالاتر به نگرش شما از نظر میزان خطر پذیری بستگی دارد.

گاهی اوقات شباهت عملکرد خویشاوندان به دلایل دیگر است. به عنوان مثال، برادران و خواهران تنی یک مادر دارند، بنابراین در معرض یک محیط مشترک، در این مورد، محیط مشترک مادری، قرار می‌گیرند. از آنجا که از نظر هر دو والد با هم مشابه هستند، لذا نسبت به حیوانات غیر خویشاوند و حتی برادر و خواهران ناتنی، ترکیبات ژنی مشترک بیشتری دارند و این منبع تشابه ژنتیکی به طور معمول به عنوان بخشی از "محیط مشترک" برادر و خواهران تنی نیز در نظر گرفته می‌شوند. به علت محیط مشترک، شباهت رکوردهای عملکرد برادر و خواهران تنی، اغلب نسبت به آنچه که به دلیل ژن‌های مشترک آنها (۵۰٪) انتظار داریم، بیشتر است. بنابراین رکوردهای بیشتر در آنها نسبت به حالت‌های دیگر اطلاعات کمتری در اختیار ما قرار می‌دهد.

از لحاظ تئوری، اثرهای محیط مشترک می‌تواند در داخل هر نوع روابط خانوادگی وجود داشته باشد. به عنوان مثال، در خانواده‌های برادر و خواهران ناتنی پدري در حیوانات شیری نیز این اثر وجود دارد. نمونه بارز اثر محیط مشترک در داخل خانواده‌های برادر و خواهران تنی وجود دارد. زمانی که اثر محیط مشترک وجود داشته باشد ولی در پیش بینی ژنتیکی در نظر گرفته نشود، ضرایب تابعیت (عوامل تصحیح) و صحت‌های مربوط به آن بیشتر از مقدار صحیح آن برآورد خواهند شد.

عوامل موثر در صحت پیش بینی ارزش اصلاحی

عواملی مانند تعداد رکوردها، وراثت پذیری، تکرار پذیری و ارتباط شجره‌ای در صحت پیش بینی می‌باشد. افزایش وراثت پذیری (بدون توجه به منبع اطلاعات)، سبب افزایش صحت پیش بینی خواهد شد. زیرا وراثت پذیری، میزان قوی بودن ارتباط بین ارزش‌های اصلاحی و ارزش‌های فنوتیپی است. هرچه این ارتباط قوی‌تر باشد، رکورد عملکرد هر حیوان شاخص بهتری از ارزش اصلاحی اوست. زمانیکه منبع اطلاعات، رکورد خود حیوان باشد، وراثت پذیری، میزان قوی بودن ارتباط بین ارزش اصلاحی مورد پیش بینی و منبع اطلاعات خواهد بود. وراثت پذیری به عنوان صحت است. زمانی که منبع اطلاعات خود فرد نباشد و برخی از خویشاوندان وی (به عنوان مثال، برادر و خواهران ناتنی) باشد، میزان وراثت پذیری بالا باعث می‌شود که رکورد عملکرد هر برادر و خواهر تنی یا ناتنی شاخص بهتری از ارزش اصلاحی خود حیوان و بنابراین شاخص بهتری از ارزش اصلاحی برادر و خواهران او (فردی که ارزش اصلاحی آن مد نظر است) باشد. همچنین با استفاده از ارتباط شجره‌ای میزان صحت افزایش می‌یابد. اعداد مربوط به ستون ارتباط شجره‌ای در جدول زیر نسبت ژن‌های مشترک در انواع مختلف خویشاوندان را نشان می‌دهد. هر فرد در ۱۰۰٪ ژن‌ها با خودش، ۲۵٪ ژن‌ها با برادر و خواهران ناتنی خود و ۵۰٪ با ژن‌های نتاج خود مشترک است. ردیف رکوردهای خود فرد را با ردیف‌هایی که شامل یک رکورد برادر و خواهران ناتنی و یک رکورد نتاج هستند، مقایسه کنید. مقدار صحت در حالت استفاده از رکورد خود فرد، بالاترین بوده و استفاده از رکورد نتاج و رکورد برادر و خواهران ناتنی به ترتیب بعد از آن قرار دارند. هر چه خویشاوندی فرد با حیوانات دارای رکورد نزدیک‌تر باشد، رکوردهای آنها شاخص بهتری از ارزش اصلاحی فرد خواهد بود.

جدول ۲- صحت پیش بینی ارزش اصلاحی با استفاده از یک منبع اطلاعات (منبع: کتاب مفاهیم اصلاح نژاد دام)

وراثت پذیری			تعداد رکوردها	شجره	ارتباط	منبع اطلاعات
۰/۷۰	۰/۳۰	۰/۰۵				ای
۰/۸۴	۰/۵۵	۰/۲۲	۱	۱	۱	خود فرد
۰/۲۱	۰/۱۴	۰/۰۶	۱	۰/۲۵		خواهر و برادران تنی
۰/۴۱	۰/۳۳	۰/۱۷	۱۰			

۰/۴۵	۰/۳۹	۰/۲۲	۲۰		
۰/۴۹	۰/۴۷	۰/۳۷	۱۰۰		
۰/۴۹ ⁺	۰/۴۹ ⁺	۰/۴۸	۱۰۰۰		
۰/۴۲	۰/۲۷	۰/۱۱	۱	۰/۵۰	نتاج
۰/۸۲	۰/۶۷	۰/۳۴	۱۰		
۰/۹۰	۰/۷۹	۰/۴۵	۲۰		
۰/۹۸	۰/۹۴	۰/۷۵	۱۰۰		
۰/۹۹ ⁺	۰/۹۹	۰/۹۶	۱۰۰۰		

با افزایش تعداد رکوردها، صحت پیش بینی افزایش می‌یابد. به ردیف‌های مربوط به برادر و خواهران ناتنی یا نتاج در جدول ۲ توجه کنید. هر چه تعداد رکوردها بیشتر باشد، صحت پیش بینی بالاتر می‌باشد. بدیهی است که با افزایش تعداد رکوردها، پیش بینی مورد نظر حاوی اطلاعات بیشتری خواهد بود.

اطلاعات موجود در جدول ۲ نشان می‌دهد که ارزش رکورد خود فرد نسبت به رکوردهای خویشاوندان متفاوت است. در مورد صفاتی با وراثت پذیری بالا، اهمیت رکورد خود فرد بالاتر است. در وراثت پذیری ۰/۷ رکورد خود فرد حاوی اطلاعات بیشتری نسبت به ۱۰ رکورد نتاج فرد و بهتر از ۱۰۰ رکورد برادر و خواهران ناتنی فرد است، که به طور مستقیم از تعریف وراثت پذیری (میزان قوی بودن ارتباط بین عملکرد و ارزش‌های اصلاحی) ناشی می‌شود. زمانی که وراثت پذیری بالا است، رکورد عملکرد خود فرد باید شاخص خوبی از ارزش اصلاحی او باشد. همان طور که در سطوح بالای وراثت پذیری، رکورد فرد از اهمیت بیشتری برخوردار است، بنابراین به صورت منطقی انتظار می‌رود که در سطوح پایین وراثت پذیری نیز، رکوردهای عملکرد خویشاوندان، با ارزش‌تر باشند. در واقع نیز به همین صورت است. در وراثت پذیری ۰/۵، رکورد عملکرد خود فرد با ۲۰ رکورد برادر و خواهر ناتنی و ۴ تا ۵ رکورد نتاج از نظر فراهم نمودن اطلاعات مشابه است. اگر تعداد رکورد خویشاوندان زیاد باشد، رکورد خود فرد نمی‌تواند اطلاعات معادل آنها را فراهم نماید.

رکوردهای فرزندان، آخرین منبع اطلاعات برای پیش بینی ارزش اصلاحی هستند. توجه داشته باشید در جدول ۲- صحت‌های بالای پیش بینی زمانی به دست می‌آیند که تعداد فرزندان زیاد باشد. با تعداد فرزندان کافی، بنا به دلایل زیر حتی در مورد صفاتی با وراثت پذیری پایین نیز، صحت بالا است. رکوردهای فرزندان به عنوان

معیاری از ارزش ژنتیکی منتقل شده از فرد می‌باشد. یک رکورد از هر نتاج بنا به دلایل مختلف ممکن است اطلاعات چندانی ارائه ندهد. (۱) اثرهای محیطی ممکن است بیشتر از اثرهای ژنتیکی باشد (وراثت پذیری پایین باشد)، (۲) ارزش اصلاحی والد دیگر در محاسبات در نظر گرفته نشده باشد، (۳) نتاج ممکن است (در اثر نمونه گیری مندلی) نمونه‌ای از ژن‌های بسیار خوب یا بد را از والد دریافت کرده باشد. اما اثرهای محیطی، ارزش‌های اصلاحی والدین و اثرهای نمونه‌گیری مندلی را اغلب می‌توان با استفاده از تعداد فرزندان زیاد برطرف نمود. میانگین عملکرد تعداد زیادی از نتاج شاخص خوبی از ارزش اصلاحی فرد خواهد بود.

پرورش دهندگان اغلب فرض می‌کنند که امکان پیشرفت ژنتیکی در صفات با وراثت پذیری پایین، به دلیل مشکل بودن شناسایی حیواناتی که از لحاظ ژنتیکی ممتاز هستند، در عمل غیر ممکن است. این مطلب زمانی درست است که منابع اطلاعات خود را به رکوردهای عملکرد حیوان محدود نماییم. اما با استفاده از رکوردهای نتاج (و تعداد زیادی از آنها) می‌توان مشکل پایین بودن وراثت پذیری را برطرف کرد. می‌توان حیوانات دارای ارزش‌های اصلاحی بهتر را به طور صحیح شناسایی نمود.

رکوردهای حاصل از برادر و خواهران تنی و ناتنی تنها تا حدودی می‌تواند منجر به افزایش صحت پیش بینی شود. در جدول ۲- به ردیف مربوط به ۱۰۰۰ رکورد از برادر و خواهران ناتنی توجه کنید. حتی با وجود اطلاعات زیاد (۱۰۰۰ رکورد) نیز، صحت پیش بینی از ۰/۵ فراتر نخواهد رفت. در حقیقت بدون توجه به تعداد رکوردها یا وراثت پذیری صفت، حد نهایی صحت پیش بینی حاصل از رکورد برادر و خواهران ناتنی برابر ۰/۵ است. این امر به این دلیل است که اطلاعات برادر و خواهران ناتنی تنها نیمی از اطلاعات شجره فرد را فراهم می‌کند و نمونه گیری مندلی در محاسبات در نظر گرفته نمی‌شود.

پیش بینی ارزش اصلاحی با استفاده از چندین منبع اطلاعات

در بسیاری موارد می‌خواهیم که ارزش اصلاحی یک فرد را بر اساس بیش از یک منبع اطلاعات پیش بینی کنیم. به عنوان مثال، حالت زیر را در نظر بگیرید:

مثال چهار معادله‌ای

فرض کنید می‌خواهیم ارزش اصلاحی حیوانی را برای یک صفت غیر تکراری بر اساس چهار منبع اطلاعات: (۱) عملکرد فرد (IND)، (۲) میانگین عملکرد برادر و خواهران ناتنی پدری (PHS)، (۳) میانگین عملکرد برادر و خواهران ناتنی مادری (MHS)، و (۴) میانگین عملکرد نتاج (PROG) پیش بینی کنیم. بنابراین شاخص انتخاب به صورت زیر خواهد بود:

$$I = b_1x_1 + b_2x_2 + b_3x_3$$

یا

$$EBV = b_1P_{IND} + b_2\bar{P}_{PHS} + b_3\bar{P}_{MHS} + b_4\bar{P}_{PROG}$$

مقایسه حیوانات با استفاده از اطلاعات حاصل از گروه‌های ژنتیکی متفاوت (بهترین پیش بینی نا

اریب خطی)

شاخص انتخاب روش قدرتمندی برای پیش بینی ژنتیکی است. اما فرض پایه شاخص انتخاب این است که اطلاعات عملکرد مورد استفاده، از گروه‌های ژنتیکی مشابه هم عصر به دست آمده است. اگر بخواهیم برای پیش بینی از اطلاعات حاصل از گروه‌های هم عصر که از لحاظ ژنتیکی متفاوت هستند (گروه‌های مزارع و مراتع یا دهه‌های مختلف) استفاده نماییم چگونه خواهد بود؟ روش توسعه یافته شاخص انتخاب که اصطلاحاً بهترین پیش بینی نااریب خطی یا BLUP نامیده می‌شود، برای استفاده در چنین شرایطی طراحی شده است.

قدمت تئوری آماری بهترین پیش بینی نااریب خطی تقریباً هم زمان با تئوری شاخص انتخاب ارائه شده است. اما تا اوایل دهه ۱۹۸۰، کامپیوتر و الگوریتم‌های ریاضی آن قدر پیشرفت نکرده بود که از BLUP به طور وسیع برای ارزیابی ژنتیکی استفاده شود. BLUP به محاسبات پیچیده نیاز دارد. همانند شاخص انتخاب که از منابع چندگانه اطلاعاتی استفاده می‌کند، BLUP نیز شامل معادلات چند مجهولی است. با وجود این، به طور معمول BLUP چندین معادله شاخص انتخاب را با هم ترکیب می‌کند، که تا حدودی به این دلیل است که یک آنالیز منفرد BLUP پیش بینی‌ها را برای همه حیوانات یک جمعیت فراهم می‌کند - نه فقط برای یک حیوان در یک زمان، در حالی که اغلب برنامه‌های شاخص انتخاب این گونه نیست. به علت توانایی BLUP در محاسبه تفاوت‌های ژنتیکی بین گروه‌های هم عصر و نیز به علت پیش بینی ژنتیکی هم زمان برای تعداد زیادی از حیوانات، به عنوان روشی برای ارزیابی ژنتیکی گسترده یعنی ارزیابی ژنتیکی در جمعیت‌های بزرگ (به عنوان مثال، کل نژادها) می‌باشد. رکوردهای عملکرد مورد استفاده برای چنین ارزیابی‌هایی به طور معمول از داده‌های مزرعه‌ای به دست می‌آید که این داده‌ها به طور منظم توسط پرورش دهندگان به انجمن‌های اصلاحی یا سازمان‌های دولتی گزارش می‌شود.

بهترین پیش بینی نااریب خطی، تکنیکی است که می‌توان آن را به عنوان نوعی مدل‌های آماری در نظر گرفت. این روش به صورت نمایش ریاضی عملکرد حیوانات است که شامل اثرات محیطی و ژنتیکی متنوع بوده و جهت پیش بینی ژنتیکی استفاده می‌شود. انواعی از مدل‌های پدری، مدل پدربزرگ پدری-مادری، مدل حیوانی، مدل اندازه‌گیری‌های تکراری، مدل مادری مستقیم، مدل چند صفتی، و غیره وجود دارد. این مدل‌ها در درجه اول از

نظر نوع حیوان مورد نظر برای پیش بینی ژنتیکی (به عنوان مثال تنها پدر، همه والدین یا همه حیوانات)، تعداد و انواع پیش بینی‌های حاصل و مشکلات محاسباتی با هم متفاوت هستند. به طور معمول هرچه مدل پیچیده باشد، دارای معادلات بیشتری بوده و به محاسبات کامپیوتری بیشتری نیاز خواهد داشت.

مجدداً یادآوری می‌کنیم که شاخص انتخاب نمی‌تواند این تفاوت را در نظر بگیرد. زیرا شاخص انتخاب، انحراف از میانگین گروه‌های هم عصر را به کار می‌برد ولی مدل‌های BLUP، انحراف از میانگین‌های گروه هم عصر را به کار نمی‌برند، بلکه در واقع شامل معادلاتی هستند که اثرهای گروه هم عصر و اثرهای محیطی مشترک (E_{cg}) برای همه حیوانات گروه هم عصر را حل می‌کنند. این مدل‌ها، این کار را با مقایسه عملکرد خویشاوندان در گروه‌های هم عصر مختلف انجام می‌دهند. اگر حیوانات در یک گروه هم عصر خاص، به خوبی عملکرد تعیین شده خویشاوندان آنها در گروه‌های هم عصر دیگر نباشند، بنابراین اثر محیط مشترک برای آن گروه هم عصر کمتر از میانگین برآورد شده است. همچنین اگر حیوانات در یک گروه هم عصر عملکردی بهتر از میزان مورد انتظار نشان دهند، E_{cg} بهتر از میانگین برآورد شده است. از طریق حل معادلات چند مجهوله، برآورد اثرهای گروه هم عصر اطلاعاتی فراهم می‌کند که برای پیش بینی ژنتیکی با قابلیت اعتماد بیشتری برای تک تک حیوانات مورد استفاده قرار می‌گیرد و برعکس. بنابراین روش‌های BLUP، در حقیقت برای پیش بینی ارزش‌های اصلاحی از اطلاعات عملکرد حاصل از گروه‌های هم عصر متفاوت از لحاظ محیطی و ژنتیکی استفاده می‌کنند و این تفاوت‌ها را در پیش بینی در نظر می‌گیرند. پیش بینی‌های ژنتیکی حاصل از مدل‌های BLUP را می‌توان به صورت مستقیم با یکدیگر مقایسه نمود، حتی اگر این پیش بینی‌ها با استفاده از رکوردهای حاصل از محیط‌های بسیار متفاوت و (یا) گروه‌های هم عصر با میانگین شایستگی ژنتیکی متفاوت صورت گرفته باشند.

روش‌های سنتی شاخص انتخاب، از رکوردهای حاصل از تعداد نسبتاً محدودی از منابع استفاده می‌کند ولی مدل‌های BLUP می‌توانند از اطلاعات تعداد زیادی از منابع استفاده کنند. در حقیقت، این مدل‌ها اغلب اطلاعات همه حیوانات جمعیت را مورد استفاده قرار می‌دهند. به این علت که رکوردهای خویشاوندان دور احتمالاً سهم چندانی در پیش بینی ندارند. با وجود این، قابلیت روش‌های BLUP در ترکیب اطلاعات انواع خویشاوندان باعث افزایش صحت پیش بینی (حتی به مقدار ناچیز) خواهد شد. همچنین نشان دهنده ارتباط شجره‌ای است. در BLUP استفاده از همه داده‌ها را می‌توان به صورت استفاده از داده‌های صفات همبسته تفسیر نمود. مدل‌های BLUP چند صفتی (Multiple-trait BLUP models) مدل‌های آماری مورد استفاده برای تخمین ارزش‌ها برای چندین صفت در زمان واحد هستند. همچنین از اطلاعات یک صفت برای پیش بینی ارزش صفت دیگر استفاده می‌شود. این زمانی مفید است که اطلاعات در مورد صفت دوم بسیار کم باشد. به

عنوان مثال، به طور معمول در داده‌های مزرعه‌ای وزن شیرگیری نسبت به وزن تولد بیشتر است. زمانی که وزن شیرگیری و وزن تولد با هم در مدل BLUP چند صفتی در نظر گرفته می‌شود، اطلاعات وزن شیرگیری به پیش بینی مقادیر وزن تولد و برعکس کمک می‌کند. در نتیجه صحت پیش بینی هر دو صفت بالاتر می‌رود.

برخی دیگر از قابلیت‌های BLUP

Non-random mating

– در نظر گرفتن آمیزش غیر تصادفی

پرورش دهندگان گله مولد تمایلی به آمیزش تصادفی حیوانات (یعنی در نظر گرفتن نرها برای آمیزش با ماده‌ها به صورت تصادفی) ندارند. آیا می‌توان آنها را مقصر دانست؟ اگر برای تلقیح گاو ماده گله خود معروف خود، ۵۰۰۰ دلار برای خرید گاو نر پرداخت کرده باشید، آیا آن را با هر گاو ماده‌ای آمیزش می‌دهید؟ البته پاسخ منفی است. با آمیزش این گاو نر با بهترین گاو ماده گله، قصد داریم احتمال تولید گوساله ممتاز را حداکثر نماییم. مشکل این راهبرد از دیدگاه پیش بینی‌ژنتیکی در نظر گرفتن برتری برخی حیوانات بر برخی دیگر است (که باعث ایجاد اریب می‌شود). در این حالت بهترین مولدین نر با بهترین مولدین ماده آمیزش داده می‌شوند، بنابراین احتمالاً دارای بهترین نتاج خواهند شد. بقیه مولدین نر نیز برای مولدین ماده معمولی‌تر در نظر گرفته شده و بنابراین احتمال دارد که نتاج با عملکرد ضعیف‌تری تولید کنند. روش‌های BLUP در اصل، اثر آمیزش غیر تصادفی را به وسیله تصحیح پیش بینی‌های حیوانات برای شایستگی جفت‌های آنها در نظر می‌گیرد. امکان پیش بینی‌های ژنتیکی مولدین نر و ماده به طور هم زمان را فراهم نموده، و حل معادلات چند مجهوله باعث می‌شود که هر پیش بینی، پیش بینی‌های دیگر را نیز در نظر بگیرد.

Poor performance culling

– حذف به خاطر عملکرد ضعیف

یکی از موارد دیگری که باعث اریب می‌شود، حذف به خاطر عملکرد ضعیف در صفت تکرارپذیر یا صفاتی است که قبل از رکورد برداری یک صفت همبسته از نظر ژنتیکی، رکورد برداری شده است. به عنوان مثال، اسب‌های تروبرد به طور معمول در سن دو سالگی مسابقه حرفه‌ای خود را آغاز می‌کنند. اگر عملکرد آنها در فصل اول مسابقات موفقیت آمیز باشد، به مسابقه ادامه خواهند داد. در غیر این صورت به دلیل عملکرد ضعیف حذف خواهند شد. در صورتی که نتاج برخی از پدران بیشتر از نتاج پدران دیگر باشد، چنین حذف‌هایی باعث اریب خواهد شد. مدل‌های BLUP چند صفتی، اریب حاصل از حذف ناشی از عملکرد ضعیف را در نظر می‌گیرند. اگر صفتی که حذف بر اساس آن انجام گرفته، و نیز رابطه ژنتیکی (همبستگی ژنتیکی) بین این صفت و صفت بعدی در آنالیز در نظر گرفته شود، بنابراین پیش بینی‌ها برای صفت دوم تحت تاثیر عمل حذف قرار نخواهند گرفت.

– امکان پیش بینی مولفه های ژنتیکی مستقیم و مادری صفات و تفکیک آنها

همه صفات دارای مولفه مستقیم (Direct component) هستند که شامل اثر ژن های فرد بر عملکردش می باشد. بعضی صفات دارای مولفه مادری (Maternal component) هستند که شامل اثر ژن های مادر فرد است که از طریق محیط فراهم شده توسط مادر بر عملکرد فرد موثر است. بهترین مثال برای صفت دارای هر دو مولفه مستقیم و مادری، وزن شیرگیری است. وزن شیرگیری حیوان تابعی از توانایی ژنتیکی قابل توارث نرخ رشد، تولید شیر و قابلیت مادری والد ماده است. توانایی ژنتیکی قابل توارث نرخ رشد، به وسیله ژن های خود حیوان تعیین می شود که شامل مولفه مستقیم وزن شیرگیری است. میزان تولید شیر و قابلیت مادری والد ماده به وسیله ژن های مادر (و همچنین توسط محیط) تعیین می شود. ژن های والد ماده برای این صفات به طور مستقیم بر نرخ رشد نتاج موثر نیستند، اما بر محیطی که نتاج در آن قرار گرفته، موثر خواهند بود. تولید شیر و قابلیت مادری مولفه های مادری وزن شیرگیری هستند.

صفات دیگری که مولفه های مادری برای آنها اهمیت دارند، شامل سخت زایی و قابلیت زنده ماندن است. مولفه مستقیم سخت زایی مربوط به اندازه و شکل جنین است. همچنین مولفه مادری به اندازه و شکل لگن والد ماده و نیز عوامل روان شناسی و فیزیولوژیکی نامحسوس تر بستگی دارد. مولفه مستقیم قابلیت زنده ماندن تابعی از ژن های موجود در حیوانات جوانی است که بر سلامت جسمی، پاسخ ایمنی و غریزه زنده ماندن تاثیر می گذارند. مولفه مادری این صفت به قابلیت مادر برای تغذیه و مواظبت از نتاج خویش بستگی دارد. روش های BLUP قابلیت تفکیک مولفه های مستقیم و مادری یک صفت را داشته و برای هر دو مولفه، پیش بینی ژنتیکی انجام می دهد.

ارزش دیگری که به طور معمول به وسیله روش BLUP پیش بینی می شود، ارزش مادری کل (BV_{tm}) Total Maternal Value نامیده می شود. ارزش مادری کل برای فرد، ترکیبی از ارزش های اصلاحی اجزای مادری و مستقیم یک صفت است. به عنوان مثال، ارزش مادری کل برای وزن شیرگیری نشانگر قابلیت ژنتیکی او برای تولید وزن شیرگیری است. به همین صورت، شامل اثر ژن های مادر برای تولید شیر و قابلیت مادری (مولفه مادری وزن شیر گیری)، و نیز اثر ژن های منتقل شده از مادر به گوساله برای نرخ رشد قبل از شیر گیری – اثر مستقیم وزن شیر گیری است. به عنوان مثال ارزش مادری کل برای وزن شیرگیری به شکل معادله به صورت زیر است:

$$BV_{WW_{tm}} = BV_{WW_m} + \frac{1}{2} BV_{WW_d}$$

در این فرمول اندیس های WW، m و d به ترتیب نشان دهنده وزن شیرگیری، مولفه مادری و مولفه مستقیم می باشد. ارزش مادری کل در گاو شامل همه ارزش اصلاحی او برای مولفه مادری وزن شیر گیری و تنها نصف ارزش اصلاحی وی برای مولفه مستقیم وزن شیرگیری است. دلیل این امر آن است که همه ژن های مادر که در تولید شیر و قابلیت مادری موثرند، بر عملکرد گوساله نیز تاثیر دارند، اما به علت انتقال تنها نیمی از ژن های مادر به گوساله، فقط نصف ژن های مادری موثر در نرخ رشد قبل از شیرگیری، عملکرد گوساله را تحت تاثیر قرار می دهد. به صورت پیش بینی ژنتیکی داریم:

$$EBV_{WW_{tm}} = EBV_{WW_m} + \frac{1}{2}EBV_{WW_d}$$

و

$$EPD_{WW_{tm}} = EPD_{WW_m} + \frac{1}{2}EPD_{WW_d}$$

به طور معمول EBV مادری کل و EPD مادری کل با هم اشتباه گرفته می شوند. در یک حیوان ماده، EBV مادری کل، شامل پیش بینی قابلیت ژنتیکی او برای تولید صفتی است که در نتاج آن اندازه گیری می شود. به عبارت دیگر، همان تولید مورد انتظار خود مادر (بدون در نظر گرفتن هر گونه اثر محیطی) است. به عبارت دیگر، EPD مادری کل یک فرد، تولید مورد انتظار دختران اوست.

بحث مولفه های مستقیم و مادری با ذکر مولفه سومی به نام مولفه پدری کامل می شود. یکی از تعاریف مولفه پدری همانند تعریف مولفه مادری است. بر اساس این تعریف، مولفه پدری یک صفت، اثر ژن های پدر بر عملکرد فرد است که از طریق محیط فراهم شده توسط پدر بر عملکرد او تاثیر گذار خواهد بود. مولفه پدری بر اساس این تعریف در مورد حیوانات اهلی کمتر صدق می کند چون پدران در پرورش نتاج نقش کمی داشته و در نتیجه هیچ اثر محیطی بر عملکرد فرزندان شان ندارند. با وجود این، مولفه پدری در مورد پرندگان وجود دارد. تعریف دوم مولفه پدری برای صفات باروری به کار می رود. معیارهای باروری که به عنوان صفات مولد ماده یا نتاج در نظر گرفته می شوند، تحت تاثیر صفات باروری و قابلیت فیزیکی حیوان نر آمیزش یافته قرار می گیرند که به آن مولفه پدری گفته می شود. مثال این تعریف، مولفه پدری نرخ آبستنی است. تعریف دوم بیشترین کاربرد عملی را در اصلاح نژاد گونه های اهلی دارد.