

DAMIZLIK DEĞERİN "EN İYİ SAPMASIZ TAHMİNLEYİCİSİ" BLUP YÖNTEMİ

Yavuz AKBAŞ

GİRİŞ

Hayvan ıslahında ilk basamak üzerinde çalışılan popülasyonun tanınmasıdır. Doğru ve düzenli olarak tutulan kayıtlar bu tanıma işlemini büyük ölçüde sağlar. İkinci aşamada tutulan kayıtları kullanarak hayvanların damızlık (genetik) değerlerin saptanması gelir. Popülasyondaki bireylere ait damızlık değerlerin doğru olarak saptanamaması halinde elde edilecek genetik ilerleme yavaş ve etkisiz olur.

Genetik ıslah çalışmasının başarısı,

- a. popülasyonda mevcut genetik varyasyonun düzeyine,
- b. genetik değerlendirmenin doğru yapılıp yapılmamasına,
- c. uygulanan seleksiyonun yoğunluğuna ve
- d. generasyon aralığına bağlıdır.

Genetik ıslahın hızını artırmada en büyük fırsatı bu faktörlerden ikincisi (değerlendirmenin doğruluğu) bize tanımaktadır. Çünkü hayvan ıslahçısı popülasyondaki genetik varyasyon ile generasyon aralığını kontrol etmede göreceli olarak daha az bir etkiye sahiptir. Genetik ıslahta seleksiyon yoğunluğu ancak çok sayıda hayvan ile çalışılması ve hayvanların genetik değerlerinin doğru bir şekilde tahmin edilmesi durumunda etkili olabilir. Ayrıca genetik değerlendirmelerin doğruluğu, diğer faktörlere göre kontrol altına alınması en kolay olanıdır.

* E. Ü. Ziraat Fakültesi Zootekni Bölümü, İzmir

Bu çalışmada genetik değerin belirlenmesinde mevcut yöntemler arasında en doğru tahminlemeyi veren "*En İyi Doğrusal Sapmasız Tahminleme*" yöntemi BLUP tanıtılmış ve yöntemin baba değerlendirmede kullanımına ait bir örnek sunulmuştur.

YÖNTEMİN TANITIMI

BLUP, C.R. Henderson (1949, 1973) tarafından seleksiyon indeksi yerine önerilmiş bir yöntemdir. Yaklaşım olarak seleksiyon indeksinin tüm olumlu istatistik özelliklerine sahiptir. Bilindiği gibi seleksiyon indeksinde verimler önce bilinen tüm çevre etmenlerince düzeltilir. Bu amaçla farklı yöntemler kullanılmaktadır (Gönül, 1974). Kimi durumlarda mevcut düzeltme faktörlerinden yararlanılırken, kimi durumlarda ise veri setinden düzeltme faktörlerinin hesaplanması gerekir. Faktör sayısının fazla olması halinde veri setinden düzeltme faktörlerinin hesaplanması ve düzeltme işleminin gerçekleştirilmesi oldukça fazla hesaplama gerektirir. İkinci aşamada indeks oluşturulup bireylere ait damızlık değerleri tahminlenir. Halbuki BLUP yöntemi bu hesaplamaların hepsini tek bir aşamada kombine eder ve sabit etkiler için daha önce herhangi bir düzeltme işlemi gerektirmez.

BLUP ismi "**Best Linear Unbiased Prediction - En İyi Doğrusal Sapmasız Tahminleme**" isminin baş harflerinden oluşmaktadır. Bu kelimeler yöntemin istatistiksel özelliklerini açıklamaktadır. **En iyi** terimi tahminlenen sonuç ile gerçek damızlık değeri arasındaki farkın minimum bir varyasyona sahip olduğunu açıklarken, diğer tanımlamalarda tahminleyicinin gözlemlerin **doğrusal** bir fonksiyonu olduğunu, tahminlenmek istenen değerin **sapmasız** tahminlendiği ve yapılan işlemin gerçek damızlık değerini **tahminleme** olduğunu açıklamaktadır.

BLUP yöntemi seleksiyon indeksinin istenen özelliklerine sahip olması yanısıra daha bir çok üstünlüklere de sahiptir. Farklı gruplandırılan bireylerin BLUP ile tahminlenen damızlık değerleri birbirleriyle doğrudan karşılaştırılabilir bir yapıdadır (Örneğin farklı yıllarda doğan bireylerin karşılaştırılması gibi). Ayrıca BLUP rastgele olmayan çiftleştirmeler, zamana bağlı olarak çevrenin değişimi, anaların damızlık değerlerinde sürü farklılıkları ile seleksiyon ve ayıklama sonucu olabilecek sapmaları da dikkate alarak tahminlemeleri yapar (Nicholas, 1987). Bu yöntemde değerlendirmeye alınan hayvanlar arasındaki akrabalık

ilişkileri hesaplamalara dahil edilebilmektedir. Bunun sonucu damızlık değerlerin tahmin doğrulukları artmaktadır (Swalze ve Bruns, 1985).

BLUP yöntemi üstün özellikleri nedeniyle tüm dünyada hayvanların damızlık değerlerinin tahminlenmesinde kullanılan yegane yöntem olmuştur. Sadece sığırcılıkta değil diğer türlerde de kullanımı yaygınlaşmıştır. Bu çalışmada yöntemin tanıtımı babaların değerlendirilmesi yaklaşımında ele alınmıştır.

UYGULAMA (Baba Değerlendirme)

Babanın eklemeli genetik değerini tahminlemede yavrularının kayıtları kullanılmaktadır (Yavru testi). Özellikle dişilerde ölçülen özellikler bakımından babaların genetik potansiyelini belirlemede yavru testi çok kullanışlı bir yöntemdir. Yöntemin esası, yavrunun, ebeveynlerden her birinin genetik potansiyelinin rastgele bir yarısına sahip olduğu genetik yaklaşımına dayanır ($G_{yavru} = 0.5 G_{baba} + 0.5 G_{ana} + \text{hata}$). Anaların popülasyondan alınan rastgele örnekler olduğunun kabul edilmesiyle ana etkisinin artan yavru sayısı ile her baba için eşit olduğu kabul edilebilir. Eşitlikle yeralan hata terimi, artan n sayısına bağlı olarak ortalaması sıfır olan rastgele hataları açıklamaktadır.

Örneğimizdeki iki sürüde kullanılan üç babanın yavrularına ait verim toplamları Tablo 1'de verilmiştir. Bu veri setini kullanarak BLUP yönteminin uygulamasını basamak basamak inceleyelim.

Tablo 1. Yavru sayısı ve verim toplamlarının sürü ve babalara göre durumu

Baba No	Sürü-1		Sürü-2		TOPLAM	
	Yavru sayısı	Toplam	Yavru sayısı	Toplam	Yavru sayısı	Toplam
1	4	480	6	840	10	1320
2	50	6250	0	0	50	6250
3	50	7250	50	6750	100	14000
TOPLAM	104	13980	56	7590	160	21570

1. BLUP öncelikle gözlemleri açıklayan modelin yazılmasını gerektirir. Modelin doğruluğu yöntemin temelini oluşturur. Çünkü eksik veya

hatalı modele BLUP uygulaması yerine, doğru modele diğer yöntemlerin uygulanması daha başarılı sonuç verir. Örnek verilerimizi açıklamak için aşağıdaki model oluşturulmuştur.

$$Y_{ijk} = m + s_i + b_j + e_{ijk}$$

Modelde m özelliğe ait genel ortalamayı, s_i , i nolu sürü etkisini, b_j , j nolu baba etkisini, e_{ijk} ise hata terimini açıklamaktadır. Buna göre Y_{ijk} i nolu sürüde, j nolu babaya ait k nolu bireyin incelenen özellik için verim değerini açıklamaktadır.

2. Modeldeki bütün etkiler sıralanıp genel toplamları bulunur.

Etki	m	s1	s2	b1	b2	b3
Toplam	21570	13980	7590	1320	6250	14000

3. Her etkinin toplamı model içindeki parametreler bakımından ayrıntılı olarak yazılır. Örneğin bir nolu sürüde 104 bireye ait ortalama etkisi, 104 tane sürü - 1 etkisi, 4 tane bir nolu baba etkisi, 50 tane iki nolu baba etkisi, 50 tane üç nolu baba etkisi varken doğal olarak bu sürüde iki nolu sürü etkisi yoktur (sıfırdır).

Bu şekilde aşağıdaki tablo oluşturulur.

Modeldeki Etkiler	m	s1	s2	b1	b2	b3	Toplam
Ortalama için	60	104	56	10	50	100	= 21570
sürü - 1 için	104	104	0	4	50	50	= 13980
sürü - 2 için	56	0	56	6	0	50	= 7590
1 nolu baba için	10	4	6	10	0	0	= 1320
2 nolu baba için	50	50	0	0	50	0	= 6250
3 nolu baba için	100	50	50	0	0	100	= 14000

Bu tablo bize "En küçük kareler (EKK)" eşitliklerini vermektedir. Bundan sonraki basamaklar matrisler kullanılarak anlatılacaktır. Buna göre EKK eşitlikleri :

160	104	56	10	50	100	m		21570
104	104	0	4	50	50	s1		13980
56	0	56	6	0	50	s2		7590
10	4	6	10	0	0	b1	=	1320
50	50	0	0	50	0	b2		6250
100	50	50	0	0	100	b3		14000

4. Baba toplamlarının olduğu satırların (son üç satır) baba etkilerine ait sütunlarına $k = \sigma^2 e / \sigma^2 b$ terimi eklenir. Bu terim hata varyansının babaya ait varyansa oranıdır. Yavru testinde baba varyansı genetik varyansın 1/4'üne eşittir.

$$\text{Kalıtım derecesi } (h^2) = 4 \sigma^2 b / (\sigma^2 b + \sigma^2 e) \text{ olur.}$$

$$h^2 (\sigma^2 b + \sigma^2 e) = 4 \sigma^2 b \quad h^2 \sigma^2 b + h^2 \sigma^2 e = 4 \sigma^2 b$$

$$h^2 \sigma^2 e = 4 \sigma^2 b - h^2 \sigma^2 b \quad h^2 \sigma^2 e = (4 - h^2) \sigma^2 b$$

Buradan $k = \sigma^2 e / \sigma^2 b$ değerinin $(4 - h^2) / h^2$ olduğu anlaşılır. Örneğimizde kalıtım derecesinin 0.25 olduğunu kabul edersek, $k = (4 - 0.25) / 0.25 = 15$ olur.

EKK eşitliklerine k katsayısının eklenmesi ile "Karışık Model Eşitlikleri" elde edilir. Bu durumda karışık model eşitlikleri aşağıdaki şekilde yazılır.

160	104	56	10	50	100	m		21570
104	104	0	4	50	50	s1		13980
56	0	56	6	0	50	s2		7590
10	4	6	10+15	0	0	b1	=	1320
50	50	0	0	50+15	0	b2		6250
100	50	50	0	0	100+15	b3		14000

5. Karışık model eşitliklerinin çözümü yapılarak modeldeki etkiler tahminlenir. Fakat hayvancılıkta kullanılan verilerin yapısı sonucu karışık model eşitliklerinin katsayılar matrisinde satır ve sütunlar arası bağımlılık nedeniyle matrisin tersini alma işlemi yapılamaz. Bu durumda bazı kısıtlamalar uygulanarak bu bağımlılık yok edilir ve ters alma işlemi yapılarak çözümler elde edilir. Örneğimizde Satır 1= Satır2 + Satır3 bağımlılığı vardır. Bu durumu yok etmek için $s1 = 0$ kısıtlaması uygularsak, x gibi hayali bir etki için yeni bir eşitlik (yeni bir satır ve sütun) daha oluşturulur. Eşitliklerimiz,

160	104	56	10	50	100	0	m	21570
104	104	0	4	50	50	1	s1	13980
56	0	56	6	0	50	0	s2	7590
10	4	6	25	0	0	0	b1	= 1320
50	50	0	0	65	0	0	b2	6250
100	50	50	0	0	115	0	b3	14000
0	1	0	0	0	0	0	x	0

olur. Eşitliği kısaca

$$Bc = T$$

şekilde de yazabiliriz. Çözüm vektörü $c = B^{-1} T$ işlemi sonucu elde edilir. Modelimizdeki etkiler için

$$\begin{aligned} m &= 134.5413 & b1 &= 0.2672 \\ s1 &= 0 & b2 &= -7.3394 \\ s2 &= -5.3487 & b3 &= 7.0722 \end{aligned}$$

çözümleri bulunur.

BLUP sonuçlarında doğrudan karşılaştırmalar yapabiliriz. Örneğin ortalama sürü etkisi altında bir nolu babanın kızlarının ortalama performansı genel ortalamaya göre 0.2672 birimlik bir üstünlüğe sahipken, iki nolu babanın kızlarına göre 7.6066 birimlik bir üstünlüğe, üç nolu babanın kızlarına göre ise 6.805 birimlik daha düşük bir verime sahiptir. Yavru testinin genetik mekanizması nedeniyle babaların tahminlenen

damızlık değerleri (estimated breeding value) hesaplanan baba etkilerinin (b'lerin) iki katıdır. Bu nedenle b değerlerine kalıtımla aktarılan kısım (transmitting ability) adı verilir. Bu durumda babaların tahminlenen damızlık değerleri baba numaralarına göre sırasıyla 0.5344, -14.6788 ve 14.1444 olur.

BLUP yöntemiyle damızlık değer tahminlerinde akrabalık ilişkileri de dikkate alınabilmektedir. Örneğimizde babalar arasında akrabalık olmadığı kabul edilmiştir. Şimdi aynı örnekte bir ve iki nolu babaların üvey kardeş olduklarını kabul edelim ($a_{12} = a_{21} = 0.25$). Bu durumda babalar arası akrabalık matrisi (A),

$$A = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = \begin{vmatrix} 1 & 0.25 & 0 \\ 0.25 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix}$$

şeklinde yazılabilir.

BLUP işlemlerinin 4. basamağında EKK eşitliklerini karışık model eşitliklerine çevirmede babalara ait alt matrisin köşegen elemanlarına k değeri eklenirken, akrabalık ilişkileri dikkate alındığında aynı alt matrise $k \cdot A^{-1}$ değeri eklenir. Burada A^{-1} akrabalık matrisinin tersini açıklamaktadır. A matrisinin tersi,

$$A^{-1} = \begin{vmatrix} 1 & 0.25 & 0 \\ 0.25 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix}^{-1} = \begin{vmatrix} 1.067 & -0.267 & 0.000 \\ -0.267 & 1.067 & 0.000 \\ 0.000 & 0.000 & 0.000 \end{vmatrix}$$

ve

$$kA^{-1} = \begin{vmatrix} 16 & -4 & 0 \\ -4 & 16 & 0 \\ 0 & 0 & 15 \end{vmatrix}$$

matrisi hesaplanır. Daha sonra EKK matrisine ilave edilip karışık model eşitlikleri oluşturulur.

160	104	56	10	50	100	m		21570
104	104	0	4	50	50	s1		13980
56	0	56	6	0	50	s2		7590
10	4	6	10 +16	-4	0	b1	=	1320
50	50	0	-4	50+16	0	b2		6250
100	50	50	0	0	100+15	b3		14000

Karışık model eşitlikleri çözümlenerek modeldeki etkiler tahminlenir. Bu durumda,

$$\begin{aligned}
 m &= 134.789 & b1 &= -1.020 \\
 s1 &= 0 & b2 &= -7.478 \\
 s2 &= -5.214 & b3 &= 6.789
 \end{aligned}$$

çözümleri elde edilir. Babaların her iki durumda tahminlenen damızlık değerleri arasındaki fark, akrabalık ilişkilerinin popülasyondaki düzeyine bağlı olarak değişir.

Baba No	babalar akraba değil	baba1 ile baba2 üvey kardeş
1	0.5344	-2.040
2	-14.6788	-14.956
3	14.1444	13.578

Baba1 ve baba2'nin akraba olması demek, baba1'in baba 2'ye, baba2'nin de baba1'in damızlık değerinin hesaplanmasında ek bilgi kaynağı oluşturduğu anlamına gelmektedir. Örneğimizde ilk iki babanın damızlık değerleri genel ortalamanın altında seyrettiği için akrabalık ilişkilerinin modele alınması ortalamalarının doğal olarak daha da düşmesine yol açmıştır.

TARTIŞMA

Görüldüğü gibi BLUP yöntemiyle damızlık değerlerin tahminlenmesinde yoğun hesaplamalar kullanılmaktadır. Hayvancılıkta kullanılan

veri setlerinin fazla sayıda olması yapılacak hesaplamaların yükünü daha da artırmaktadır. Fakat hızlı gelişen bilgisayar teknolojisi ve buna bağlı olarak uygun yazılımların geliştirilmesi söz konusu zorlukların azalmasını sağlamıştır. Bununla birlikte bilgisayar kullanarak matrislerin doğrudan tersinin alınmasında da matrisin boyutlarına bağlı olarak zorluklarla karşılaşmaktadır. Henderson (1975), BLUP için akrabalık matrisinin tersinin alınmasında doğrudan ve kolay çözümlene getiren bir yaklaşım sunmuştur.

Bütün anılan zorluklara rağmen sürülerde uygulanacak seleksiyon kararlarında güvenilir bilgi vermesi nedeniyle BLUP tüm Dünyada yaygın olarak kullanılmaya başlanmıştır. Çünkü ıslah başarısını doğrudan etkileyen ve ıslahçının diğer faktörlere göre en kolay kontrol edebildiği nokta, genetik değerlendirmelerin doğru ve güvenilir yapılp yapılmamasıdır.

BLUP yöntemi örneğimizdeki gibi babaların için (baba modeli) veya popülasyondaki her birey (örneğin bütün inekler) için uygulanabilir (birey modeli). Uygulanacak yaklaşım genel olarak aynıdır. Sadece uygun k değerinin hesaplanması ve her bireyin modele bir etki olarak alınması gereklidir. Baba modelinde babaların mümkün olduğu kadar farklı sürüde çok sayıda kızının olması damızlık değerini tahminleme gücünü arttırmaktadır. Birey modeli kullanımının bize getireceği zorluk, popülasyonda birey sayısının babaların sayısından fazla olması ve matrislerin boyutlarının artmasından kaynaklanmaktadır. Örneğin sığırcılıkta sürülerde kullanılan az sayıdaki babaya karşılık çok sayıda inek bulunmaktadır.

BLUP yöntemi birden fazla özelliği dikkate alan modellere ve/veya farklı genetik komponentleri (örneğin eklemeli olmayan genetik etkileri) içeren modellere kolaylıkla uygulanabilir.

KAYNAKLAR

1. Gönül, T. 1974. Hayvan ıslahında standardizasyon, Tavukçuluk Araştırma Enstitüsü, TABGE yayınları No: 15, Ankara.
2. Henderson, C. R. 1949. Estimation of chages in herd environment. Journal of Dairy Science 32 : 706.

3. Henderson, C. R. 1973. Sire evaluation and genetic trends. In animal Breeding and Genetics. (Proceedings of a Symposium in Honor of Dr. J. L. Lush.) (Ed. Anon.) pp. 10-41. American Society of Animal Science, and American Dairy Science Association, Champaign, Illinois.
4. Henderson, C. R. 1976. A simple method for computing the inverse of a numerator relationship matrix used in prediction of breeding values. *Biometrics* 32 : 69-83.
5. Nicholas, F. W. 1987. *Veterinary Genetics*. Clarendon Press, Oxford.
6. Swalve, H., Bruns, E. 1985 . A note on the properties of a relationship matrix in mixed model sire evaluation. *Z. Tierzüchtg. Züchtgsbiol.* 102 : 331-341.