

Kümeleme Analizi ile Seleksiyon

İsmet DOĞAN

Afyon Kocatepe Üniversitesi, Veteriner Fakültesi, Biyometri Anabilim Dalı, Afyon - TÜRKİYE

Geliş Tarihi: 06.07.2000

Özet: Bu çalışmanın amacı, farklı disiplinlerde kullanım yeri bulmuş olan Kümeleme Analizi Yöntemi'nin hayvan ıslahı ile ilgili çalışmalarda da kullanılabileceğini göstermektir.

Çalışmanın materyalini Çifteler Tarım İşletmesinde yetiştirilen ve 1980-1998 yılları arasında satışa sunulan 535'i erkek, 392'si dişi olan toplam 927 Arap tayı oluşturmıştır.

Kümeleme metodu olarak k-Ortalama Yöntemi kullanılmıştır. Kümelerin belirlenmesinde yaş, cidago yüksekliği, göğüs çevresi ve incik çevresine ait değerler dikkate alınmıştır. Kullanılan bilgiler, satışa sunulan atlar ile ilgili olarak hazırlanan satış kataloglarından derlenmiştir.

Kümelere her bir cinsiyet için ayrı ayrı elde edilmiştir. Erkek taylar 16, dişi taylar ise 14 farklı kümeye ayrılmıştır.

İstatistiksel analizlerin uygulanmasında SPSS for Windows Release 10.0 istatistik paket programından yararlanılmıştır.

Anahtar Sözcükler: Kümeleme Analizi, k-Ortalama Yöntemi, Hayvan Islahı, Seleksiyon, Arap tay

Selection by Cluster Analysis

Abstract: The aim of this study was to show that cluster analysis, which has been used in different disciplines, can be used in animal breeding studies.

This study included 535 Arabian colts and 392 Arabian fillies (total: 927) that were offered for sale between 1980 and 1998, bred at Çifteler State Farm.

As a cluster method, the k-means method was used. In terms of determination of clusters, the data used were age, height at withers, heart girth and cannon bone circumference. This information was obtained from the marketing catalogue.

The clusters were obtained for each sex separately. Colts were divided into 16 clusters, and fillies were divided into 14 clusters.

For the application of statistical analysis, SPSS for Windows Release 10.0 was used.

Key Words: Cluster Analysis, k-Means Method, Animal Breeding, Selection, Arabian Colts/Fillies

Giriş

Hayvan ıslahı, üzerinde çalışılan populasyonun (sürünün) belirli bir veya birkaç özellik (verim) bakımından genotipik değerini şimdiki nazaranda daha yüksek seviyelere çıkarmak amacıyla yapılan faaliyetler olarak tanımlanmaktadır (1). Mevcut populasyonun ıslahına, üzerinde durulan verim / verimler bakımından varyasyonun tespiti ve analizi ile başlanır. Bu analiz neticesinde verimli veya istenen özelliklere sahip olan

bireyler damızlıkta kullanılmak üzere alıkonulur. Böylece bir grup içindeki bazı bireylere diğerlerine göre daha fazla döl verme olanağı sağlanır. Bu işleme "seleksiyon" denir (2, 3). Yirminci yüzyılın ilk başlarına kadar seleksiyon genellikle fiziki uygunluğa ve ırk karakterlerine göre yapılmaktaydı. Ancak pedigrî kayıtları tutulmaya başlandıktan sonra ferdi kayıtlara ve ferdin ailesinin kayıtlarına bakılarak seçim yapılabilmesi için metotlar geliştirilmiştir (3).

Seleksiyon yöntemleri genel olarak;

1. Bireylerin kendi fenotipik değerlerine göre,
 2. Akrabaların fenotipik değerlerine göre,
 - a. Pedigriye göre,
 - b. Yavru ortalamalarına göre (progeny testing),
 3. Baraj metoduna göre,
 4. İndeks metoduna göre,
- şeklinde sınıflandırılmaktadır (2, 4-7).

Yukarıda verilen yöntemlerin hemen tamamında gözleme dayalı yada kayıtlardan yararlanarak veri elde edilmesi söz konusudur. Verilerin toplanmasında kullanılan bu metotlar bilimsel metotlar olarak kabul edildikleri için, bu metotların uygulanmaları sonucunda elde edilen sonuçlar da bilimsel değerleri olan sonuçlar olarak değerlendirilebilir. Bu metotlar bilimsel olarak kabul edilseler dahi bunların da bazı zayıf yönleri bulunmaktadır. Örneğin, fenotipe göre yapılan seçim işlemi, hızlı sonuç vermesine rağmen genotipi tam olarak yansıtmadığı için neredeyse terk edilmiş durumdadır. Pedigri yada yavru ortalamalarına göre yapılan seçim işleminde ise çok sayıda kayıt tutulması gerekmektedir. Baraj metoduna göre yapılacak seçimde ise barajın ne olacağı konusunun belirlenmesinde kullanılan herhangi bir objektif metot bulunmamaktadır. Seleksiyon işleminde en etkili metot olan indeks metodunda ise değerlendirmeye alınan hayvanların en son kayıtları tamamlanıncaya kadar elde bulundurulmaları gerekmektedir. Bu da büyük zaman alıcı önemli bir faktördür (2).

Verilen bu metotların ortak noktası bir sürü içindeki bütün bireylerin, dikkate alınan özellik yada özellikler bakımından sahip oldukları sayısal değerlerin büyüklükleri bakımından birbirleriyle karşılaştırılıyor olmalarıdır. Dolayısıyla herhangi bir özellik bakımından çok yüksek değer alan bir birey diğer özellikler bakımından çok kötü değerler de olsa çok yüksek değer aldığı özelliğinden dolayı diğer bireylere nazaran tercih edilebilecektir. Bu ise tek bir özelliğinden dolayı, seçilmemesi gereken bir bireyin seçilmesini sağlamaktadır ki bu bütün metotların en büyük zaafiyetidir.

Bilimsel araştırmalarda, araştırmaya konu olan olaylar veya nesnelere her birey için aynı anda ölçülebilen bir veya birden çok değişken tarafından etkilenebilirler. Birden çok değişkenin ayrı ayrı ele alınarak analiz edilmesi, gerçek

durumu açıklamayabilir. Çünkü, değişkenlerin ayrı ayrı ele alınarak analiz edilmesi, değişkenler arasındaki ilişkileri dikkate almamak demektir. Oysa gözlemlenen bu çok sayıda değişken arasında az veya çok bir ilişkinin olması beklenmektedir. Bu amaçla "Çok Değişkenli İstatistiksel Analiz Yöntemleri" geliştirilmiştir. Çok değişkenli istatistiksel analiz, çok sayıda değişken arasındaki ilişkileri ölçme ve açıklamada kullanılan yöntemler topluluğunu ifade eder ve bu analiz ile ilgili olarak geliştirilmiş yöntemler, bağımlılık analizinde kullanılan yöntemler ve karşılıklı bağımlılık analizinde kullanılan yöntemler olmak üzere iki grupta toplanabilir (8).

Bağımlılık analizinde, bir değişken veya değişken grubu, diğer değişkenler tarafından tahmin edilmekte veya açıklanmaktadır. Burada, bir değişken diğerlerine bağlı olup onlarla tahmin edilmektedir. Karşılıklı bağımlılık analizinde ise değişkenler arasındaki bağımlılık yerine karşılıklı bağıntılar söz konusudur. Bu tip analizin en iyi örneklerinden biri Kümeleme Analizi' dir (8).

Kümeleme Analizi özellikle son yıllarda popüler olan, çok değişkenli istatistiksel yöntemlerden biridir. Bu yöntem, özellikle bilim ve iş alanında, bir çok durumda uygulanabilen, en etkili ve en kolay yorumlanabilen bir yöntem olma özelliği taşımaktadır (9-12). Kümeleme Analizi'nin genel amacı, gruplandırılacak verileri, benzerliklerine göre alt sınıflara ayırarak açıklamaktır. Başka bir ifade ile, çalışmada yer alan tüm değişkenler itibari ile bireyler veya nesnelere arasındaki benzerlikler esas alınarak, benzer bireylerin aynı gruplarda veya kümelerde toplanması ve yeni bir bireyin hangi gruba dahil olduğunun tahmin edilmesi Kümeleme Analizi'nin esasını teşkil etmektedir (13, 14).

Kümeleme Analizi, bireylerin yada nesnelere sınıflandırılmasını ayrıntılı bir şekilde açıklamak amacıyla geliştirilmiştir. Bu amaca yönelik olarak, ele alınan örnekte yer alan varlıklar aralarındaki benzerliklere göre gruplara ayrılır, daha sonra bu gruplara dahil edilen bireylerin profili ortaya konur. Bir başka ifade ile bu analizde, özellikle amaçlanan şey, öncelikle ele alınan örnekte gerçekte var olduğu bilinen grupları, varlıklar (birey yada nesne) arasındaki benzerliklere dayanan az sayıdaki karşılıklı özel grupları oluşturmak, daha sonra bu gruplara giren varlıkların profilini ortaya koymaktır (15).

Kümelemede pek çok yöntem bulunmakta ve bu yöntemler farklı başlıklar altında toplanmaktadır. Kümeleme yöntemleri, birim ya da değişkenleri uygun

gruplara ayırırken grupları belirlemede izledikleri yaklaşımlara göre;

1. Aşamalı Kümeleme Yöntemleri (Hierarchical Cluster Analysis Methods).

2. Aşamalı Olmayan Kümeleme Yöntemleri (Nonhierarchical Cluster Analysis Methods) biçiminde iki temel gruba ayrılmaktadır (9, 13, 14, 16).

Aşamalı kümeleme yöntemlerinde birimlerin birbirleri ile birleştirilmesinde değişik yaklaşımlar uygulanmaktadır. Bu yöntemlerden sıklıkla kullanılan ve bilinenler;

1. Tek Bağlantı Kümeleme Yöntemi,
2. Ortalama Bağlantı Kümeleme Yöntemi,
3. Tam Bağlantı Kümeleme Yöntemi,
4. Mc Quitty Bağlantı Kümeleme Yöntemi,
5. Küresel Ortalama Bağlantı Kümeleme Yöntemi,
6. Ortanca Bağlantı Kümeleme Yöntemi,
7. Ward Bağlantı Kümeleme Yöntemi'dir.

Aşamalı olmayan kümeleme yöntemlerinin gerek teorik dayanaklarının aşamalı kümeleme yöntemlerine göre daha güçlü olması gerekse küme sayısı konusunda ön bilgi olması ya da araştırmacının anlamlı olacak küme sayısına karar verebilmesi aşamalı kümeleme yöntemlerine tercih edilmesini sağlamaktadır.

Aşamalı olmayan kümeleme yöntemleri arasında en çok kullanılan yöntemler;

1. k-Ortalama Yöntemi,
2. En Çok Olabilirlik Yöntemi'dir.

Kümeleme Analizi'nde, orjinal değerler ya da standartlaştırılmış değerler kullanılabilir (15). Ayrıca önerilen bu metotta, bireylerin dikkate alınan özelliklerine ait bireysel değerlerinin büyüklükleri değil, bütün bireylerin incelenen tüm özellikler bakımından birbirlerine göre uzaklıkları dikkate alındığından birey seçmek yerine dikkate alınan bütün özellikler bakımından gruplar oluşturmak söz konusudur. Dolayısıyla fert seçiminden ziyade grup seçimi söz konusu olduğundan zaman kaybı önlenmekte ve tüm özellikler bakımından toplu değerlendirmeye müsaade edildiğinden diğer yöntemlere üstünlük sağlamaktadır.

Kümeleme Analizi'nde, küme sayısının belirlenmesi konusunda son yıllarda yoğun çalışmalar yapılmaktadır. Halen küme sayısının belirlenmesinde kullanılan en pratik yol, n kümelenecek birey sayısını göstermek üzere;

$$k = \sqrt{\frac{n}{2}} \quad (1)$$

biçiminde belirtilmektedir (14).

Küme, birbirlerine yakın bireylerin çok boyutlu uzayda oluşturdukları birlik olarak ifade edilebilir. Bu durumda küme kavramı, "benzerlik" ve "uzaklık" kavramlarını çağrıştırmaktadır. Noktaların geometrik olarak gösterimlerinde ikiden fazla boyut olduğunda noktalar arasındaki uzaklıkları çok boyutlu olarak hesaplamak gerekir. P sayıda değişkene göre birimler yada değişkenler arasındaki uzaklıkları hesaplamak için en sık kullanılan uzaklık ölçüleri;

1. Minkowski Uzaklığı,
2. Öklid (Euclide) Uzaklığı,
 - a. Ölçekli Öklid Uzaklığı,
 - b. Binary Öklid Uzaklığı,
3. Pearson Uzaklığı,
4. Manhattan (City-Blok) Uzaklığı,
5. Mahalanobis Uzaklığı,
6. Hotelling T² Uzaklığı,
7. Canberra Uzaklığı,

olarak verilmektedir (12, 16, 17). Kümeleme analizinde veri girişinden sonra, hesaplanan uzaklık değerlerinden yararlanarak birey ya da nesnelerin kümelere atanması işlemi yapılmaktadır. Kümeleme yöntemleri, uzaklık matrisi ya da benzerlik matrisinden yararlanarak birimler ya da değişkenleri kendi içinde homojen ve kendi aralarında heterojen gruplar oluşturmayı sağlayan yöntemlerdir. Veri matrisinde yer alan n birimin p değişkene göre uzaklıkları, uzaklık matrisi adı verilen D matrisi ile gösterilir. D matrisinin elemanları d_{ij} ya da $d(i,j)$ biçimindedir ve $n \times n$ uzaklık matrisi aşağıdaki şekilde gösterilmektedir,

$$D = \begin{bmatrix} 0 & d_{12} & d_{13} & \dots & d_{1n} \\ & 0 & d_{23} & \dots & d_{2n} \\ \text{SİM} & & 0 & \dots & \vdots \\ & & & 0 & d_{n-1,n} \\ & & & & 0 \end{bmatrix}$$

D uzaklık matrisinin uzaklık fonksiyonu olması için gerekli şartlar şunlardır; X,Y ve Z Öklid uzayında üç nokta olmak üzere,

- a) Negatif olmama, $D(X,Y) \geq 0$;
- b) Simetri, $D(X,Y) = D(Y,X)$;
- c) Teşhis işaretleri, $D(X,X) = 0$
- d) Kesinlik, $D(X,Y) = 0$ ancak ve ancak $X=Y$ ise,
- e) Üçgen eşitsizliği, $D(X,Y) \leq D(X,Z) + D(Z,Y)$

koşulları geçerli ise D , bir uzaklık olarak adlandırılır (18). Birimlerin birbirleriyle olan benzerlik düzeyleri benzerlik (S) matrisi ile gösterilir. Benzerlik matrisinin elemanları D matrisinin elemanlarına göre belirlenir. $S = 100 * [1 - d_{ij} / \max(d_{ij})]$ matrisinin elemanları S_{ij} ya da $S(i,j)$ biçiminde gösterilir. S benzerlik matrisinin gösterimi,

$$D = \begin{bmatrix} S_{11} & & & & \\ S_{21} & S_{22} & & & \\ S_{31} & S_{32} & S_{33} & \dots & \\ \vdots & \vdots & \vdots & \vdots & \\ S_{n1} & S_{n2} & S_{n3} & \dots & S_{nn} \end{bmatrix}$$

gibidir (15, 17).

Bu çalışmanın amacı, farklı disiplinlerde kullanım yeri bulunmuş olan Kümeleme Analizi Yöntemi'nin hayvan ıslahı ile ilgili çalışmalarda da kullanılabileceğini göstermektir.

Materyal ve Metot

Çalışmanın materyalini, Çifteler Tarım İşletmesinde yetiştirilen ve 1980-1998 yılları arasında satışa sunulan 535'i erkek, 392'si dişi olan toplam 927 Arap tayı oluşturmuştur.

Kümeleme metodu olarak k -Ortalama Yöntemi kullanılmıştır. Bu tekniğin bilgisayar algoritmalarındaki pratik işleyişi,

1. İlk k gözlemin her biri bir gözlemlenmiş küme olarak alınır,
 2. Kalan $n-k$ gözlemin her biri, ortalaması en yakın olan kümeye atanır ve her atamadan sonra küme ortalamaları yeniden hesaplanır,
 3. Tüm bireylerin kümelere atanması bittikten sonra, n bireyin son bulunmuş küme ortalamalarına göre yeniden atamaları yapılır,
 4. Bir önceki kümelemeye göre son elde edilen kümelemede kümeler arası birey geçişi durana kadar üçüncü adım tekrarlanır,
- şeklindedir (17).

k -ortalama yönteminde bireylerin kümelere atanmasında uzaklık ölçütü olarak Mahalanobis Uzaklığı kullanılmaktadır (15). Mahalanobis Uzaklığı, Öklid Uzaklığı'nın genelleştirilmiş bir biçimidir. $n * p$ boyutlu bir veri matrisinde birimler arasındaki Öklid Uzaklığı;

$$d(i,j) = \sum_{k=1}^p (x_{ik} - x_{jk})^2$$

$$i = 1, 2, \dots, n,$$

$$j = 1, 2, \dots, n,$$

$$n = \text{birey / birim sayısı},$$

$$p = \text{değişken sayısı},$$

x_{ik} ve y_{jk} ise sırasıyla i . ve j . bireylerin k . özelliğe ait değerler,

şeklinde hesaplanmaktadır. Mahalanobis Uzaklığı ise,

$$d^2(i,j) = (\bar{x}_i + \bar{x}_j)^t S^{-1} (\bar{x}_i - \bar{x}_j)$$

\bar{x}_i : i . bireye ait gözlem vektörü,

\bar{x}_j : j . bireye ait gözlem vektörü,

t : vektör devriği (transpozisi),

formülü ile hesaplanır. Formülde yer alan S^{-1} , benzerlik matrisinin tersini göstermektedir.

Kümelere belirlenmesinde taylara ait yaş, cidago yüksekliği, göğüs çevresi ve incik çevresine ait standartlaştırılmış değerlerden yararlanılmıştır. Standartlaştırma işlemi, her bir cinsiyet ayrı ayrı dikkate alınarak, değişkenlere ait değerler ile değişkene ait ortalama arasındaki fark, standart sapmaya bölünerek yapılmıştır. Kümeler her bir cinsiyet için ayrı ayrı elde edilmiştir. Kullanılan bilgiler, satışa sunulan taylar ile ilgili olarak hazırlanan satış kataloglarından derlenmiştir.

Kullanılan metodun gerçekten birbirinden farklı kümeler oluşturup oluşturmadığını anlamak için her bir değişken ayrı ayrı dikkate alınarak Varyans Analizi yapılmıştır. Varyans Analizi'nde erkek taylar için 4'üncü, 8'inci, ve 12'inci kümeler, dişi taylar için ise 1'inci ve 3'üncü kümeler denek sayılarının yetersizliğinden dolayı dikkate alınmamıştır.

Bulgular

Erkek taylar için k -ortalama yöntemi kullanılarak elde edilen kümelere ait sonuçlar Tablo 1'de, dişi taylara ait sonuçlar ise Tablo 2'de verilmiştir.

Tablo 1. Erkek Taylara Ait Kümeler ve Bu Kümeler İçin Değişken Değerleri

	Cidago Yüksekliği (cm) $\bar{X} \pm S_{\bar{X}}$	Göğüs Çevresi (cm) $\bar{X} \pm S_{\bar{X}}$	İncik Çevresi (cm) $\bar{X} \pm S_{\bar{X}}$	Yaş (Gün) $\bar{X} \pm S_{\bar{X}}$
1.Küme (n = 21)	145,119 ± 0,386	155,33 ± 0,60	18,8095 ± 0,0807	451,62 ± 5,83
2.Küme (n = 17)	157,765 ± 0,420	180,29 ± 0,94	19,8824 ± 0,0912	858,65 ± 10,93
3.Küme (n = 37)	138,149 ± 0,273	151,78 ± 0,57	17,8919 ± 0,0343	420,62 ± 5,32
4.Küme (n = 1)	149,0 ± 0,000	164,00 ± 0,00	17,0000 ± 0,0000	724,00 ± 0,00
5.Küme (n = 70)	142,786 ± 0,162	153,57 ± 0,36	17,8786 ± 0,0258	437,00 ± 3,88
6.Küme (n = 59)	153,678 ± 0,195	173,54 ± 0,37	19,1864 ± 0,0361	862,78 ± 6,64
7.Küme (n = 55)	150,018 ± 0,227	170,51 ± 0,42	18,9364 ± 0,0319	839,67 ± 7,19
8.Küme (n = 5)	145,000 ± 0,758	165,20 ± 1,83	18,4000 ± 0,1871	786,40 ± 14,39
9.Küme (n = 49)	150,990 ± 0,179	176,37 ± 0,31	19,7347 ± 0,0360	863,33 ± 5,30
10.Küme (n = 32)	149,297 ± 0,278	169,47 ± 0,50	19,6563 ± 0,0416	837,72 ± 9,44
11.Küme (n = 41)	154,439 ± 0,192	178,29 ± 0,43	20,1707 ± 0,0569	899,56 ± 7,72
12.Küme (n = 2)	135,000 ± 0,000	136,50 ± 3,50	16,0000 ± 0,0000	373,50 ± 3,50
13.Küme (n = 15)	139,400 ± 0,542	153,00 ± 1,04	18,7333 ± 0,0667	442,40 ± 7,70
14.Küme (n = 52)	152,321 ± 0,224	172,06 ± 0,31	20,1250 ± 0,0431	876,92 ± 8,08
15.Küme (n = 26)	146,115 ± 0,361	160,92 ± 0,66	18,3077 ± 0,0625	471,08 ± 4,30
16.Küme (n = 53)	139,849 ± 0,240	148,87 ± 0,45	17,0849 ± 0,0420	411,06 ± 3,95

Tablo 2. Dişi Taylara Ait Kümeler ve Bu Kümeler İçin Değişken Değerleri

	Cidago Yüksekliği (cm) $\bar{X} \pm S_{\bar{X}}$	Göğüs Çevresi (cm) $\bar{X} \pm S_{\bar{X}}$	İncik Çevresi (cm) $\bar{X} \pm S_{\bar{X}}$	Yaş (Gün) $\bar{X} \pm S_{\bar{X}}$
1.Küme (n = 2)	140,500 ± 2,500	146,50 ± 1,50	14,5000 ± 0,5000	393,00 ± 9,00
2.Küme (n = 8)	147,625 ± 0,754	173,87 ± 0,95	17,1875 ± 0,0915	815,75 ± 7,96
3.Küme (n = 1)	129,000 ± 0,000	136,00 ± 0,00	14,0000 ± 0,0000	384,00 ± 0,00
4.Küme (n = 19)	147,368 ± 0,427	164,58 ± 1,03	18,0526 ± 0,0362	800,84 ± 7,45
5.Küme (n = 61)	139,861 ± 0,180	153,31 ± 0,37	17,2049 ± 0,0317	428,87 ± 3,89
6.Küme (n = 16)	135,938 ± 0,421	141,87 ± 0,86	16,3438 ± 0,0880	400,31 ± 6,30
7.Küme (n = 15)	139,067 ± 0,384	152,53 ± 1,19	16,2000 ± 0,0816	423,73 ± 8,87
8.Küme (n = 17)	149,735 ± 0,515	168,35 ± 0,77	18,8824 ± 0,0530	783,65 ± 21,88
9.Küme (n = 49)	143,408 ± 0,264	158,06 ± 0,50	17,3469 ± 0,0363	453,73 ± 3,63
10.Küme (n = 78)	143,808 ± 0,198	158,68 ± 0,42	18,1090 ± 0,0268	453,03 ± 3,07
11.Küme (n = 53)	139,708 ± 0,236	152,68 ± 0,44	18,0755 ± 0,0366	434,68 ± 4,59
12.Küme (n = 19)	149,053 ± 0,451	174,89 ± 0,71	18,0526 ± 0,0362	841,11 ± 9,96
13.Küme (n = 42)	137,560 ± 0,288	148,19 ± 0,51	17,1071 ± 0,0320	409,79 ± 5,18
14.Küme (n = 12)	152,875 ± 0,564	177,50 ± 0,85	19,0000 ± 0,1231	821,83 ± 21,20

Erkek taylar için oluşturulan kümelerle ait Varyans Analizi sonuçları Tablo 3'te, dişi taylar için oluşturulan kümelerle ait Varyans Analizi sonuçları ise Tablo 4'te verilmiştir.

Varyans Analizi tablolarından da görüldüğü gibi dikkate alınan tüm değişkenler bakımından elde edilen kümeler arasında istatistiksel olarak anlamlı farklılıklar bulunmuştur.

Tablo 3. Erkek Taylar İçin Oluşturulan Kümelere Ait Varyans Analizi Sonuçları

		Kareler Toplamı	Serbestlik Derecesi	Kareler Ortalaması	F Değeri	P Değeri
Cidago Yüksekliği	Gruplar Arası	17121.15	12	1426.763	525.320	0,000
	Grup İçi	1396.018	514	2.716		
	Genel	18527.17	526			
Göğüs Çevresi	Gruplar Arası	59765.49	12	4980.457	472.087	0,000
	Grup İçi	5422.628	514	10.550		
	Genel	65188.11	526			
İncik Çevresi	Gruplar Arası	511.899	12	42.658	536.023	0,000
	Grup İçi	40.906	514	0,07958		
	Genel	552.805	526			
Yaş	Gruplar Arası	23422343	12	1951861.87	709.818	0,000
	Grup İçi	1413401	514	2749.807		
	Genel	24835744	526			

Tablo 4. Dişi Taylar İçin Oluşturulan Kümelere Ait Varyans Analizi Sonuçları

		Kareler Toplamı	Serbestlik Derecesi	Kareler Ortalaması	F Değeri	P Değeri
Cidago Yüksekliği	Gruplar Arası	6649.359	11	604.487	212.408	0,000
	Grup İçi	1072.897	377	2.846		
	Genel	7722.256	388			
Göğüs Çevresi	Gruplar Arası	25497.27	11	2317.934	172.545	0,000
	Grup İçi	5064.536	377	13.434		
	Genel	30561.80	388			
İncik Çevresi	Gruplar Arası	167.478	11	15.225	174.713	0,000
	Grup İçi	32.854	377	0,08714		
	Genel	200.332	388			
Yaş	Gruplar Arası	8827672	11	802515.597	738.834	0,000
	Grup İçi	409494.3	377	1086.192		
	Genel	9237166	388			

Tartışma

k-ortalama yöntemi kullanılarak tayların kümelere ayrıldığı bu çalışmada, dikkate alınan faktörlerin genetik yapıdan etkilendikleri düşünülerek, kümeleme işlemi her bir cinsiyet ayrı ayrı dikkate alınarak gerçekleştirilmiştir. Cidago yüksekliği bakımından erkek taylarda minimum değer 134 cm, maksimum değer 162 cm, dişi taylarda ise bu değerler sırasıyla 129 cm ve 156 cm, göğüs çevresi

bakımından erkek taylarda minimum değer 133 cm, maksimum değer 187 cm, dişi taylarda ise bu değerler sırasıyla 136 cm ve 182 cm, incik çevresi bakımından erkek taylarda minimum değer 16 cm, maksimum değer 21 cm, dişi taylarda ise bu değerler sırasıyla 14 cm ve 20 cm olarak elde edilmiştir. Bu değerlere göre genel olarak dişi tayların erkek taylara nazaran cidago yüksekliği, göğüs çevresi ve incik çevresi bakımından daha düşük

değerler aldığı söylenebilir. Tayların satışa sunuldukları andaki yaşlarına bakıldığında erkek tayların minimum 360, maksimum 1011 günlük, dişi tayların ise minimum 361, maksimum 997 günlük oldukları görülmektedir. Dolayısıyla satış işlemlerinde erkek ve dişi taylar için farklı yaşlarda satış yapılmadığı ve dikkate alınan değişkenler bakımından cinsiyetler arasında görülen farklılıkların yaş'tan ziyade cinsiyetten kaynaklanan genetik yapıya bağlı olduğu söylenebilir.

Gerek erkek taylara gerekse dişi taylara ait kümelere bakılarak üzerinde durulan özellik bakımından seçim yapılabilir. Eğer ilgilenilen özelliğe göre yüksek değerlere sahip bireylerin seçilmesi isteniyorsa erkek taylarda cidago yüksekliği ve göğüs çevresi için ikinci, incik çevresi için on birinci ya da on dördüncü kümenin, dişi taylar için ise her üç özellik bakımından da on dördüncü kümenin seçilmesi gerekir. Eğer dikkate alınan değişkenlere ait en düşük değerler bakımından seçim yapılması isteniyorsa, erkek taylar için cidago yüksekliği bakımından üçüncü kümenin, göğüs çevresi bakımından on ikinci kümenin, incik çevresi bakımından ise ya on ikinci kümenin ya da on

altıncı kümenin seçilmesi gerekir. Dişi taylar için ise, her üç değişkene göre üçüncü kümenin seçilmesi gerekir.

Görüldüğü gibi Kümeleme Analizi'nde, dikkate alınan iki özellik bakımından herhangi bir kümenin tercih edilmesine rağmen üçüncü özellik bakımından bir başka kümenin tercih edilmesi ya da her bir değişken için ayrı ayrı kümelerin tercih edilebilir olması gibi sonuçlar ile karşılaşılması mümkündür. Bu tür sonuçların ortaya çıkması Kümeleme Analizi'nde bir ödünleşim problemini gündeme getirmektedir. Yani iki özellik bakımından tercih edilen bir kümenin üçüncü bir özellik bakımından tercih edilememesi ya da her bir değişken için farklı kümelerin seçilmesinin gerektiği durumlarda karar vericinin bir karar vermesi gerekebilir. Bu tür sonuçların ortaya çıktığı durumlarda karar vericinin, dikkate alınan değişkenlerin önemine göre küme seçmesi gerekir.

Diğer seleksiyon metotlarına göre zaman, işgücü, maliyet, güvenilirlik, objektiflik vb. yönlerden avantajlı olmasından dolayı hayvan ıslahında özellikle de seleksiyon yaparken bir metot olarak Kümeleme Analizi'nin kullanılmasının uygun olacağı söylenebilir.

Kaynaklar

1. Düzgüneş, O., Eliçin, A., Akman, N., Hayvan Islahı, Ankara Üniversitesi, Ziraat Fakültesi Yayınları, No: 1437, Ankara, 1996.
2. Arıtürk, E., Genel Zootečni, Fırat Üniversitesi, Veteriner Fakültesi Yayınları, No: 14, Ankara Üniversitesi Basımevi, Ankara, 1977.
3. Arıtürk, E., Aksoy, T., Tavuklarda Ekonomik Karakterlerin Kalıtımı ve Kantitatif Seleksiyon, Ankara Üniversitesi Veteriner Fakültesi Yayınları, No: 312, 1975, Ankara.
4. Alpan, O., Sığır Yetiştiriciliği ve Besiciliği, 2.Basım, Ankara, 1992
5. Arıtürk, E., Hayvancılıkta Uygulanan Yetiştirme ve Islah Yöntemleri, Ankara Üniversitesi, Veteriner Fakültesi Yayınları, No: 335, Ankara, 1975.
6. Arıtürk, E., Yalçın, B. C., Hayvan Yetiştirmede Seleksiyon, Ankara Üniversitesi Veteriner Fakültesi Yayınları, No: 194, Ankara, 1966.
7. Yalçın, B. C., Genel Zootečni, İstanbul Üniversitesi Veteriner Fakültesi Yayınları, İstanbul, 1981.
8. Çakır, F., Karşılıklı Bağlılığın Ölçülmesinde Kümeleme Analizi ve Bir Uygulama, Marmara Üniversitesi Sosyal Bilimler Enstitüsü (Yüksek Lisans tezi), İstanbul, 1994.
9. Anderberg, M. R., Cluster Analysis for Application, Academic Press, New York, 1973.
10. Kachigan, S. K., Multivariate Statistical Analysis, Radius Press, New York, 1982.
11. Krzanowski, W. J., Principles of Multivariate Analysis A User's Perspective, Clarendon Press, 1988, Oxford.
12. Press, S. J., Applied Multivariate Analysis, Holt, Rinehart and Winston Inc., Chicago, 1972.
13. Hair, J. F., Anderson, R. E., Tatham, R. L., Black, W. C., Multivariate Data Analysis with readings, Prentice-Hall International Inc., Fourth Edition, New Jersey, 1992.
14. Tatlıdil, H., Uygulamalı Çok Değişkenli Analiz, Ankara, 1996.
15. Kurt, G., Çok Değişkenli İstatistiksel Analiz Tekniklerinden Çok Boyutlu Ölçekleme ve Bir Uygulama, Marmara Üniversitesi Sosyal Bilimler Enstitüsü (Doktora Tezi), İstanbul, 1992.
16. Özdamar, K., Paket Programlar ile İstatistiksel Veri Analizi II (Çokdeğişkenli İstatistiksel Yöntemler), Osman Gazi Üniversitesi Tıp Fakültesi, Eskişehir, 1998.
17. Hand, D. J., Discrimination and Classification, John Wiley & Sons, Londo, 1986.
18. Gürkanlı, A. T., Genel Topoloji, Ondokuz Mayıs Üniversitesi Yayınları, Yayın No: 73, Samsun, 1993.