TARIM BiLiMLERi DERGIS[2004, 10 (1) 67-71
DOI: 10.1501/Tarimbil_0000000871

Kanonik Korelasyon Analizi ve Bir Uygulaması

Sıddık KESKIN[1] Abdullah Nuri ÖZSOY[1]

Geliş Tarihi: 20.01.2003

Özet : l<anonik korelasyon analiz], her birinde en az iki değişken bulunan, k1 değişken seti arasındaki ilişkilerfn incelenmesinde kullanılan bir analiz tekniğidir. Bu anallıde, değişken sellerinden birisi açıklayıcı yada bağımsız değişken seti, diğeri ise bağımlı değişken seti olarak tanımlanabilir. Ancak, değişken setlerinin bu şekilde tanımlanması zorunluluğu yoktur. Bu analizde, değişken setleri arasındaki korelasyonun maksimum olması amaçlanır ve bu amaca yönefik, her iki değişken selinde yer alan değişkenlerin doğrusal kombinasyonlarından yent değişken (kanonlk değişken} çiftleri elde edilir. Bu çalışmada, kanonik korelasyon anallzf tanıtılmış ve konunun anlaşılmasını k.olaylaştırmal< amacıyla, hayvancılık alanından bir uygulama yapılmıştır.

Anahtar Kellmeler: kanonik korelasyon, kanonik değişken, değişken seti, redundancy indeksi

Canonical Correlation Analysis and Its an Application

Abstract : Canoriical correlation analysis is employed to study relationships between two variabte sets when each variable set consist of at least two variables. in this analysis, it may be defined as one set of varables is the predfctor or Tndependent set and another set of variables is the dependenl variable set. However, it is not oecessary to define the set of variables like ttiis. in !his analysts, it is aimed that the relatlonships between sets of variables is maximum. For thls aim, new varlable pafrs (caninical variables) are obtained from the flnear combfnations of the varlables in each set in this study, canonicat correlation analysls was descrlbed by lhe apptication in the cattle deatıng field to facilitale comprehensive of the matter.

Key Words : cııonical corelation, canonical varlable, varlable set, redundancy lndex

Giriş

Bilimsel çalışmalarda genellikle, aynı zamanda birderı faz.la özellik üzerinde durulmaktadır. Konu, farklı uygulama yada muamele gruplarının bir biri ile karşılaştırılması olduğunda, bu özelliklerin teker teker ele alınıp incelenmesi uygun bir yaklaşım olacaktır. Ancak, bunların bırlikte dikkate alınarak, uyğun analiz yöntemleıi ile değerlendirilmesinin, araştırıcıya ek bilgiler sağlayacağı da unutulmamalıdır. Benzer durum özellikler arasındaki doğrusal ilişkinin araştırılmasında da geçerlidir. ôzellikler arasındaki doğrusal ilişkiyi belirlemede ilk akla gelen yöntem. bu özellikleri ikişer ikişer ele alarak bunlar arasında Pearson korelasyon katsayısını hesaplamaktır. Pearson lwrelasyon katsayısı. bazı varsayımlar yada ön şartlar yerine geldiğinde. sürekli değişkenler arasındaki doğrusal ilişkinin derecesinin ve yönünün belirlenmesinde en yaygın olarak kullanılan ölçüdür. Eğer ön şartlar yerine gelmemişse, Pearson korelasyon katsayısı yerine parametrik olmayan korelasyon katsayıları olan Spearman rcınk korelasyonu ve Kendal tau korelasyonu gibi korelasyon katsayıları kullanılabilir. Bazı durumlarda, özelliklerden birisi bağımlı değişken, diğeri de bağımsız. değişken olarak ele alınabilir. Bu dutumda, iki değişken arasındaki l<orelasyon katsayısı. bu değişkenler arasındaki sebep-sonuç ilişkisinin bir ölçüsü olarak da değerlendirilebilir. Bağımlı değişken bir adet, ancak bağımsız değişkenler blr den fazla ise bağımlı değişke ile bağımsız <1eğişkenler arasındaki ilişki çoklu korelasyon katsayısı ile belirlenir. Bazı durumlarda, bağımlı ve bağımsız değişkenlerin her ikisi de birden fazla olabilfr: Bu durumda, iki değişken seti arasındaki ilişkiyi belirlemede, yukarıda belirtilen korelasyon katsayılarından hiçbiri kullanılamaz. Bunların yerine, değişken setleri yada kümelerini, bu setlerde yer alan değişkenlerin doğrusal bileşenlerinden oluşan kanonik değişkenlere dönüştürerek, bu kanonik değişkenler arasındaki (doğrusal) ilişkiyi bulma temeline dayalı kanonik korelasyon kullanılır. {Gilrbüz 1989) Kanonik korelasyon analizi, çok değişkenli analiz tekniklerinden biri olup, faktör analizi ile birlikte en karmaşık işlem aşamalarını gerektiren teknikler arasında yer almaktadır (Tatlıdll 1996). Bunun yanı sıra; elde edilen sonuçların yorumlanmasındaki güçlükler, bu anaıtz tekniğinin kullanımı geri plana itmiştir. Oysa ki, biyolojik çalışmaların bir çoğunda, basit illşki katsayıları yerine kanonik korelasyon analizinin kullanılması, araştırıcıya ek bilgiler sağlayacaktır.

Materyal ve Yöntem

Bu çalışmanın materyalini; Ankara üniversitesi Ziraat Fakültesi Zootekni Bölümü Biyometri ve Genetik Anabillm Dalı Bıldırcın işletmesinde bulunan 102 adet bıldırcın ve bunların yumurtaları oluşturmaktadır. Bıldırcınlara ilişkin

[1] YO.tOncu Yıl Oniv. Ziraat Fak. Zoı:ıtekni Bölümü -Varı

Çıkış Ağırlığı (ÇA), Dördüncü Haftadaki Canlı Ağırlık (DHCA) ve Cinsi Olgunluk Ağırlığı (COA) özellikleri birinci değişken setini (X), İlk Yumurtlama Yaşı (İYY), İlk On Yumurta Ağırlığı (İOYA) ve Onuncu Hafta Yumurta Ağırlığı (OHYA) özellikleri ise ikinci değişken setini (Y) oluşturmuştur.

Kanonik korelasyon analizinin ön şartları :

1) Ele alınan özellikler bakımından verilerin çok değişkenli normal dağılım göstermesi
2) Üzerinde durulan özellikler bakımından ölçüm hatasının minimum seviyede olması
3) Ele alınan özellikler arasında çoklu bağlantı (multicolinearity) olmaması
4) Elde edilen sonuçlara güvenilirlik bakımından, örnek genişliğinin mümkün olduğunca büyük olması (Değişken sayısının 5 katı kadar).

Kanonik korelasyon ve kanonik değişkenler : p ve q >1olmak üzere; birinci değişken setinde (kümesinde) p ve ikinci değişken setinde de q adet ($q \geq p$) değişken olduğu durumda, bu iki setteki değişkenlerin doğrusal kombinasyonları alınarak, bunlar arasındaki korelasyon hesaplanabilir. Bu şekilde hesaplanan korelasyon katsayılarına kanonik korelasyon, değişkenlerin doğrusal kombinasyonlarından oluşan yeni değişkenlere de kanonik değişkenler adı verilir. Bu kanonik değişkenler ve bunlar arasındaki kanonik korelasyonlar, birbirlerinden bağımsız olacak şekilde hesaplanırlar (Kendall 1980, Johnson ve Wichern 2002).

X değişken seti $\underset{\sim}{X}' = [X_1,..., X_p]$ olarak, Y değişken seti de $\underset{\sim}{Y}' = [Y_1,..., Y_q]$ olarak gösterildiğinde: Bu değişken setlerine ait ortalama vektörü

$$\underset{\sim}{\mu} = \begin{bmatrix} \underset{\sim}{\mu_1} \\ \underset{\sim}{\mu_2} \end{bmatrix} \text{ olarak, Kovaryans matrisi de}$$

$$\Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix} \text{ olarak yazılır.}$$

Ortalama vektörü ve kovaryans matrisi örnekten hesaplandığında sırasıyla;

$$\underset{\sim}{\overline{X}} = \begin{bmatrix} \underset{\sim}{\bar{x}_1} \\ \underset{\sim}{\bar{x}_2} \end{bmatrix} \text{ ve } S = \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix} \text{ olur.}$$

$\underset{\sim}{X}$ değişken setinden hesaplanan $U = \underset{\sim}{a}'\underset{\sim}{X}$ doğrusal kombinasyonu ile $\underset{\sim}{Y}$ değişken setinden hesaplanan $V = \underset{\sim}{b}'\underset{\sim}{Y}$ doğrusal kombinasyonu arasında korelasyon hesaplanabilir. U ile V arasındaki korelasyon, bunların katları arasındaki korelasyon ile aynı olduğundan,

$$Var(U) = \underset{\sim}{a}' S_{11} \underset{\sim}{a} = 1$$

$$Var(V) = \underset{\sim}{b}' S_{22} \underset{\sim}{b} = 1$$

$$E(U) = E(\underset{\sim}{a}' \underset{\sim}{X}) = \underset{\sim}{a}' E(\underset{\sim}{X}) = 0$$

$$E(V) = E(\underset{\sim}{b}' \underset{\sim}{Y}) = \underset{\sim}{b}' E(\underset{\sim}{Y}) = 0$$

dönüşümleri yapıldığında; U (kanonik) değişkeni ile V (kanonik) değişkeni arasındaki kanonik korelasyon;

$$r_{UV} = \underset{\sim}{a}' S_{12} \underset{\sim}{b}$$

olarak hesaplanır. Bu durumda; yukarıda belirtilen şartlara bağlı olarak bu ifadenin maksimum yapılması gerekmektedir. Bu ifade, λ ve γ Lagrange çarpanları olmak üzere;

$$\varphi = \underset{\sim}{a}' S_{12} \underset{\sim}{b} - 0.5\ \lambda\ (\underset{\sim}{a}' S_{11} \underset{\sim}{a} - 1) - 0.5\ \gamma\ (\underset{\sim}{b}' S_{22}\ \underset{\sim}{b} - 1)$$

olarak Lagrange fonksiyonu biçiminde yazılabilir. Bu fonksiyonun $\underset{\sim}{a}'$ ve $\underset{\sim}{b}'$ ne göre türevi alınıp sıfıra eşitlendiğinde; elde edilen değerler yukarıdaki koşulları sağlayacaktır.

$$S_{12} \underset{\sim}{b} - \lambda S_{11} \underset{\sim}{a} = 0$$

$$S_{12} \underset{\sim}{a} - \gamma S_{22}\ \underset{\sim}{b} = 0$$

İlk eşitlik $\underset{\sim}{a}'$ ile ikinci eşitlikte $\underset{\sim}{b}'$ ile çarpılırsa;

$$\underset{\sim}{a}' S_{12} \underset{\sim}{b} - \lambda\ \underset{\sim}{a}' S_{11} \underset{\sim}{a} = 0$$

$$\underset{\sim}{b}' S_{12} \underset{\sim}{a} - \gamma\ \underset{\sim}{b}' S_{22}\ \underset{\sim}{b} = 0 \text{ olur.}$$

$$\underset{\sim}{a}' S_{11} \underset{\sim}{a} = 1 \text{ ve } \underset{\sim}{b}' S_{22}\ \underset{\sim}{b} = 1$$

olduğundan, $\lambda = \gamma = \underset{\sim}{a}' S_{12} \underset{\sim}{b}$ olur. $S_{12} = S_{21}$ olduğundan türevi alınıp sıfıra eşitlenen ifadeler ;

$$-\lambda\ S_{11} \underset{\sim}{a} + S_{12}\ \underset{\sim}{b} = 0$$

$$S_{21} \underset{\sim}{a} - \lambda\ S_{22}\ \underset{\sim}{b} = 0$$

olarak yazılır. Bu ifadeler matris gösterimi ile;

$$\begin{bmatrix} -\lambda S_{11} & S_{12} \\ S_{21} & -\lambda S_{22} \end{bmatrix} \begin{bmatrix} \underset{\sim}{a} \\ \underset{\sim}{b} \end{bmatrix} = 0$$

olarak yazılabilir. Daha önce belirtilen şartlara uygun olarak, bu matrisin çözümü için

$$\begin{vmatrix} -\lambda S_{11} & S_{12} \\ S_{21} & -\lambda S_{22} \end{vmatrix} = 0 \text{ ya da } | S_{12} S_{22}^{-1} S_{21} - \lambda^2 S_{11} | = 0 \text{ ve}$$

$$| S_{11}^{-1} S_{12} S_{22}^{-1} S_{21} - \lambda^2 I | = 0$$

olmalıdır. Bu determinant p. dereceden polinom olup, $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_p$ olacak şekilde p adet köke sahiptir. Böylece $\lambda = \underset{\sim}{a}' S_{12} \underset{\sim}{b}$; $U = \underset{\sim}{a}' X$ ve $V = \underset{\sim}{b}' Y$ arasındaki korelasyondur. En büyük korelasyon istendiğinden, $\lambda = \lambda_1$ alınır. Bu şekilde hesaplanan doğrusal kombinasyonlar yani kanonik değişkenler arasındaki korelasyonların azalan sırada oldukları unutulmamalıdır.

Kanonik korelasyon katsayılarının önem kontrolü: Kanonik korelasyon analizi, boyut indirgeme için de kullanılabileceğinden, değişken setleri arasındaki korelasyonun elde edilen yeni değişken çiftlerinden kaç tanesi ile büyük ölçüde açıklanabileceğinin diğer bir ifade ile kanonik değişken çiftleri arasındaki korelasyonlardan kaç tanesinin önemli olduğunun bilinmesi istenilebilir. Bunun için bir kaç test yöntemi önerilmiştir. Bunlardan birisi Rao (1951) tarafından önerilen F yaklaşımıdır (Thompson 1985). Bu yaklaşımda genel olarak iki değişken seti arasındaki ilişkinin önemli olup olmadığı kontrol edilir. İkinci ve daha sık kullanılan test yöntemi ise Bartlett (1941) tarafından önerilen test yöntemidir (Thompson 1985). Bu testte hesaplanan test istatistiği χ^2 (Khi-kare) istatistiğidir. Bu istatistik;

$$\chi^2 = -[n-0.5(v_1+v_2+1)] \times \log(\Lambda)$$

olarak hesaplanır. Bu eşitlikte; n: Gözlem sayısını, v_1: Birinci setteki değişken sayısını, v_2: ikinci setteki değişken sayısını göstermektedir. Lambda (Λ) ise $\Lambda = (1-R^2_{k1})(1-R^2_{k2}) \ldots (1-R^2_{kp})$ olarak bulunur. Buradan hesaplanan χ^2 istatistiği, pxq serbestlik dereceli χ^2 tablo değeri ile karşılaştırılır. Eğer H_0 hipotezi ret edilirse, en büyük olan katsayı çıkarılarak, yeniden test yapılır. Bu teste H_0 hipotezi, kabul edilene kadar devam edilir ve en son aşamada kaç tane kanonik korelasyon katsayısının önemli olduğu belirlenir.

Test (H_0) hipotezinde p adet kanonik korelasyonun sıfıra eşit olduğu, Karşıt (H_1) hipotezde en az bir adet korelasyon katsayısının sıfırdan farklı olduğu ileri sürülür.

Redundancy indeksinin hesaplanması: Büyük örneklerde, küçük kanonik korelasyon katsayıları istatistik olarak önemli olabilirken, X ve Y değişken setleri arasında hesaplanan büyük kanonik korelasyon katsayıları da bu setler arasında güçlü bir korelasyonun olduğunu belirtmeyebilir. Çünkü kanonik korelasyon, X ve Y değişkenlerinin doğrusal bileşenlerini maksimize eder. Bu nedenle, X ve Y değişken setlerinden herhangi birindeki varyasyonun diğeri tarafından açıklanan kısmını belirtmez. Bunun için, Steward ve Love (1968) tarafından önerilen redundacy (gereksizlik) indeksi hesaplanır (Sharma 1996). Bu indeks, setlerden birindeki varyasyonun diğeri ile açıklanabilen kısmını belirtir. Redundancy indeksi (RI) her kanonik korelasyon için hesaplanabilir (Sharma 1996). U_i ve V_i kanonik değişken setleri arasında hesaplanan i. kanonik korelasyon için redundancy indeksi ($RI_{U_i V_i}$) iki aşamada hesaplanır. Birinci aşamada Y değişken setindeki yada Y değişkenlerinin bulunduğu kümedeki varyasyonun i. kanonik değişken (V_i) ile ortalama açıklanabilen kısmı bulunur. Bu değer;

$$OV(Y|V_i) = \frac{\sum_{j=1}^{p} LY_{ij}^2}{p}$$

eşitliği ile hesaplanır (Sharma 1996). Bu eşitlikte, $OV(Y|V_i)$; Y değişken setindeki varyasyonun i. kanonik değişken (V_i) ile ortalama açıklanabilen kısmını ve LY_{ij}; Y değişken setindeki j. değişken ile i. kanonik değişken arasındaki yapısal korelasyonu (j. değişkenin yükünü) göstermektedir. İkinci aşamada, redundancy indeksi; $RI_{U_i V_i} = OV(Y|V_i) \times C^2_i$ (C^2_i i. kanonik korelasyonun karesidir) olarak hesaplanır.

Bir setteki varyasyonun diğer setteki değişkenler ile toplam açıklanabilen kısmı, toplam redundancy olarak adlandırılır. Bu katsayı;

$$TRI_{Y|X} = \sum_{i=1}^{p} RI_{X|Y_i}$$

olarak hesaplanır. Toplam redundancy indeksi, Y değişken setindeki varyasyonun X değişken seti ile açıklanabilen kısmını belirtir.

Çalışmada, üzerinde durulan özellikler bakımından yapılan hesaplamalar için STATISTICA for windows (ver: 5.0) istatistik paket programı kullanılmıştır (Anonymous 1995).

Bulgular ve Tartışma

Çalışmada ele alınan özelliklere ait tanıtıcı istatistikler Çizelge 1' de verilmiştir. Bu özellikler arasındaki korelasyon katsayıları ise Çizelge 2' de verilmiştir. Çizelge 2 incelendiğinde; Çıkış ağırlığı (ÇA) ile Dördüncü hafta canlı ağırlığı (DHCA), İlk yumurtlama yaşı ve Onuncu hafta yumurta ağırlığı (OHYA) özellikleri arasındaki korelasyon katsayısının istatistik olarak önemli olduğu görülür. Ayrıca; DHCA ile COA ve İYY arasındaki korelasyon katsayıları da istatistik olarak önemlidir (p<0.01). En yüksek korelasyon katsayısı, İOYA ile OHYA arasında bulunurken (0.612), COA ile İOYA ve İYY ile OHYA arasında sırasıyla; 0.351 ve -0.215 olarak hesaplanan korelasyon katsayıları da istatistik olarak önemli bulunmuştur.

Çalışmada her iki değişken setinde de eşit (p=q =3) sayıda değişken bulunduğundan, 3 adet kanonik değişken çifti ve 3 adet kanonik korelasyon elde edilmiştir. Elde edilen kanonik değişken çiftleri arasında hesaplanan kanonik korelasyonlar ve bu korelasyonlara ait standart hatalar Çizelge 3' te verilmiştir.

Çizelge 3 incelendiğinde; ilk kanonik korelasyon katsayısının % 1 düzeyinde, ikinci kanonik korelasyon katsayısının da % 5 düzeyinde önemli olduğu görülür. % 13. 3 olarak bulunan üçüncü kanonik korelasyon katsayısı ise istatistik olarak önemli değildir. Bu durumda, % 1 düzeyinde önemli olan ilk kanonik korelasyon dikkate alındığında; ilk kanonik değişken çiftine göre kanonik katsayılar ve standardize edilmiş kanonik katsayılar Çizelge 4' te verilmiştir.

Çizelge 4' teki standardize edilmemiş yani orijinal değerlerden hesaplanan katsayılar tek (unique) değildir. Yalnızca, katsayıların oranı tekdir. Bu nedenle standardize edilmiş katsayılar, orijinal değişkende meydana gelen 1 standart sapmalık artışa karşılık, kanonik değişkende standart sapma cinsinden meydana gelen değişim miktarını gösteren katsayılardır. Diğer bir ifade ile bu katsayılar, bir setteki kanonik değişkenin oluşmasında, o sette yer alan orijinal değişkenlerin etki miktarlarını (katkılarını) gösteren katsayılardır (Sharma 1996).

Buna göre; U_1 ve V_1 kanonik değişkenlerine ait eşitlik;

U_1 = - 0.073 ÇA - 0.01DHCA + 0.753 COA
V_1 = 0.978 İYY + 0.137 İOYA + 0.073 OHYA

olarak yazılır. Bu durumda; U_1 kanonik değişkenin oluşmasında; COA değişkenine ait etki miktarı (katkı) 0.753 olurken, ÇA ve DHCA değişkenlerinin katkısı negatif yönde ve COA' nın katkısına göre oldukça düşüktür. Benzer şekilde, V_1 kanonik değişkeninin oluşmasında İYY değişkeninin katkısı en büyük olmuştur. İOYA ve OHYA değişkenlerinin katkısı ise İYY' nin katkısına göre oldukça düşüktür.

Standardize edilmiş kanonik katsayılar, standardize edilmemiş kanonik katsayılara göre daha fazla tercih edilmesine rağmen, bu katsayılar da; özellikle küçük örneklerde ve veri setinde çoklu bağlantı (multicollinerarity) olduğu durumda büyük farklılıklar gösterebilmektedir. Bu nedenle; kanonik değişken ile o sette yer alan orijinal değişkenler arasındaki korelasyonların kullanılmasının daha uygun olacağı belirtilmiştir Sharma 1996). Bu korelasyon katsayıları, yükler (loadings) yada yapısal korelasyonlar (structural correlations) olarak adlandırılır. Birinci kanonik değişken çiftine ait kanonik yükler (yapısal korelasyonlar) Çizelge 5' te verilmiştir.

Çizelge 5' te X değişken setindeki değişkenlerin kendi kanonik değişkeni (U_1) ile olan kanonik yükleri incelendiğinde; en büyük yük değerinin (-0.720) DHCA değişkenine ait olduğu ve bunu -0.534 değeri ile ÇA' nın izlediği görülür. Standardize edilmiş katsayılara (Çizelge 4) göre en yüksek katkıya sahip olan COA değişkenine ait yük değeri ise en düşük olmuştur. Y değişken setinde; İYY ve OHYA değişkenlerine ait kanonik yük değeri ile standardize edilmiş katsayı (Çizelge 4) arasında önemli bir farklılık görülmezken, İOYA değişkenine ait yük değeri yaklaşık iki kat olmuştur. X değişken seti bağımsız, Y değişken seti de bağımlı değişken seti olarak kabul edildiğinde; X değişken setinin, Y değişken setine ait birinci kanonik değişkeni büyük ölçüde negatif yönde etkilediği söylenebilir. Diğer bir ifade ile ÇA ve DHCA değişkenlerindeki artışın; büyük ölçüde İYY, İOYA ve OHYA değişkenlerinden oluşan birinci kanonik değişkeninde azalmaya neden olduğu söylenebilir.

Çizelge 1. Çalışmada ele alınan özelliklere ait tanıtıcı istatistikler

Özellikler	$\bar{X} \pm S_{\bar{x}}$	Min	Mak.
ÇA (gr)	12.67 ± 2.20	8.00	17.60
DHCA (gr)	103.98 ± 15.00	69.70	143.40
COA (gr)	192.79 ± 17.63	140.00	224.70
İYY (gün)	48.76 ± 4.20	39.00	61.00
İOYA (gr)	10.68 ± 0.95	7.22	13.55
OHYA(gr)	11.58 ± 0.92	6.90	13.70

Çizelge 2. İki değişken seti için setler içi ve setler arası korelasyon katsayıları

	X değişken seti			Y değişken seti		
	ÇA	DHCA	COA	İYY	İOYA	OHYA
ÇA	1.000	0.554**	0.120	-0.408**	0.034	0.241*
DHCA		1.000	0.420**	-0.532**	0.008	0.165
COA			1.000	0.176	0.351**	0.192
İYY				1.000	0.143	0.215*
İOYA					1.000	0.612**
OHYA						1.000

* : $p< 0.05$, ** ;$p< 0.01$

Çizelge 3. Kanonik korelasyon katsayıları

Kanonik korelasyon	$r_{UV} \pm S_{r_{UV}}$
r_{UV1}	0.705** ± 0.050
r_{UV2}	0.312* ± 0.089
r_{UV3}	0.133 ± 0.097

* : $p< 0.05$, ** ;$p< 0.01$

Çizelge 4. Birinci kanonik değişken çiftine ait katsayılar

X değişken seti			Y değişken seti		
	Kanonik katsayılar	Standardize edilmiş kanonik katsayılar		Kanonik katsayılar	Standardize edilmiş kanonik katsayılar
ÇA	-0.033	-0.073	İYY	0.023	0.978
DHCA	-0.066	-0.010	İOYA	0.144	0.137
COA	0.043	0.753	OHYA	0.079	0.073

Çizelge 5. Birinci kanonik değişken çiftine ait kanonik yükler

X değişken seti			Y değişken seti		
	U_1	V_1		V_1	U_1
ÇA	-0.534	-0.376	İYY	0.982	0.692
DHCA	-0.720	-0.507	İOYA	0.321	0.226
COA	0.325	0.229	OHYA	0.053	0.038

Çizelge 6. Redundancy analizi sonuçları

Kanonik korelasyon	Y değişken seti				
	Açıklama oranı	Eklemeli açıklama oranı	Kanonik r^2	Redundancy indeksi	Eklemeli redundancy indeksi
1	0.357	0.357	0.497	0.177	0.177
2	0.538	0.895	0.097	0.052	0.229
3	0.106	1.000	0.018	0.002	0.231

Çizelge 6 incelendiğinde; Y değişken setinde, Açıklama oranı sütununda yer alan katsayılar, kanonik değişkenlere ait ortalama açıklama paylarını göstermektedir. Y değişken setindeki birinci kanonik değişken, bu setteki varyasyonun ortalama % 35.7' sini açıklarken, ikinci kanonik değişken % 53.8' ini ve üçüncü kanonik değişken de % 10.6' sını açıklamaktadır. İkinci kanonik değişkene ait ortalama açıklama payının, birinci kanonik değişkene ait ortalama açıklama payından daha yüksek oluşu, ikinci kanonik değişkenin Y setindeki varyasyonu açıklamada birinci kanonik değişkenden daha etkili olduğunu düşündürebilir. Ancak, redundancy indeksi değerleri incelendiğinde; birinci kanonik korelasyon için redundancy indeksi değeri % 17.7 olarak bulunurken, ikinci kanonik korelasyon için bu değer % 5.2 olarak bulunmuştur. Bu durumda, birinci kanonik korelasyon için % 17.7 olan redundancy indeksi, Y değişken setindeki açıklanabilen varyasyonun, % 17.7' sinin X değişken setinde yer alan değişkenler ile yada X değişken seti ile açıklanabildiğini göstermektedir. Diğer bir ifade ile birinci kanonik korelasyon adına, Y değişken setindeki varyasyonun % 17.7 si X değişken setindeki varyasyondan kaynaklanmaktadır. Çizelge 6' dan toplam redundancy indeksi hesaplandığında; $TRI_{Y|X}$ = 0.231 (0.177+0.052+0.002) olarak bulunur. Dolayısıyla, Y değişken setindeki varyasyonun % 23.1' i X değişken seti ile açıklanabilmektedir. Bunun büyük bir kısmının (yaklaşık % 76.6' sının (0.177/ 0.231)) birinci kanonik değişken çiftine ait olduğu dikkat çekmektedir.

Sonuç

Kanonik korelasyon analizi; işlem aşamalarının uzun oluşu, hesaplama adımlarının bilgisayar olmadan yapılamaması ve elde ede edilen sonuçların yorumlanmasındaki güçlükler gibi bazı nedenlerden dolayı zor bir analiz tekniği olarak düşünülebilir. Ancak günümüzde yaygın olarak kullanılan istatistik paket programları ile kanonik korelasyon analizi rahatlıkla yapılabilmektedir. Bilimsel çalışmalarda, en az iki değişkenden oluşan iki set arasındaki ilişki yapısının, bütünlüğü bozulmadan ortaya konulabilmesi ve yorumlanabilmesi bakımından kanonik korelasyon analizinin önemi yadsınamaz. Bitki ve hayvan ıslahı ile ilgili çalışmalarda; erken tespit edilebilen özellikler ile geç tespit edilebilen ve ekonomik öneme sahip olan özellikler arasındaki ilişkilerin belirlenmesi ve buna göre de ıslah çalışmalarında, erken seleksiyon yapabilme olanaklarının sağlanması bakımından bu gibi çalışmalara ihtiyaç duyulmaktadır.

Bu çalışmanın, bitki ve hayvan ıslahında, uygulamalı biyolojide ve diğer ilgili alanlarda çalışan araştırıcılara, çok değişkenli analiz yöntemlerinin ve buna bağlı olarak da kanonik korelasyon analizinin, önemini kapsamını ve uygulamada kullanılabilirliğini göstermek bakımından yararlı olabileceği ümit edilmektedir.

Kaynaklar

Anonymous, 1995, STATISTICA for windows, Relase 5.0, StatSoft Inc. USA.

Gürbüz, F. 1989, Değişken Takımları Arasındaki İlişkilerin Kanonik Korelasyon Yöntemi ile Araştırılması, Ankara Üniversitesi Ziraat Fakültesi, Yayın No. 1162, 55s., Ankara.

Johnson, R. A. and D. W. Wichern, 2002. Applied Multivariate Statistical Analysis. Prentice - Hall, Inc., Upper Saddle,762p. New Jersey.

Kendall, M. G. 1980. Multivariate Analysis. Charles Griffin & Company, LTD , 210p. London.

Sharma, S. 1996. Applied Multivariate Techniques. John Wiley & Sons, Inc. 493p. Canada.

Tatlıdil, H. 1996. Uygulamalı Çok Değişkenli İstatistiksel Analiz. Cem Web Ofset Ltd. Şti. 424s. Ankara.

Thompson, B. 1985. Canonical Correlation Analysis. Sage Publication Ltd.,69p. London.

İletişim adresi :
Sıddık KESKİN
Yüzüncü Yıl Üniversitesi Ziraat Fakültesi
Zootekni Bölümü-Van
E-mail: skeskin973@hotmail.com