

KOVARYANS ANALİZİ

Yrd.Doç.Dr. Ezel (Tavşancıl) TARKUN*

Kovaryans analizi özünde, istenmeyen değişkenlerin bağımlı değişkene yaptığı etkilerin doğrusal regresyonla dengelendiği varyans analizidir. Bu analizde önce doğrusal regresyon yapılarak istenmeyen değişkenlerin gruplara etkileri hesaplanmaktadır. Regresyon katsayıları hesaplandıktan sonra da, bu katsayılar kullanılarak bağımlı değişkenden istenmeyen değişkenlerin (kontrol değişkenlerinin) hesaplanan etkileri çıkarılmakta ve bu şekilde dengelenmiş olan verilere varyans analizi uygulanarak daha sağlıklı sonuçlar elde edilmektedir (Popham,1967).

Kovaryans analizi üç temel amaç için kullanılmaktadır(Tabachnick,Fidell, 1989; Kirk, 1968):

1. Deneysel araştırmalarda, hata terimlerinden ortak değişkenle(covariate-birlikte değişen, kontrol edilen değişken) ilgili kestirilebilir varyansları kaldırarak ana etki ve ortak etki için F testinin gücünü artırmaktadır. Bilindiği gibi deneysel araştırmalarda denemenin (işlemin) etkisini tam olarak belirleyebilmek için grupların baştan eşit olması koşulu vardır. Birçok deneysel araştırmada denekler gruplara yansız olarak atanmasına rağmen yansız atama grupları arasında hiçbir şekilde eşitliğin bir garantisi olamamaktadır. Yansız atama yalnızca gruplar arasında başlangıçta sistematik farklılıklar olmadığı garantisini belli olasılık sınırları içerisinde verir. Deneysel araştırmalarda amaç, daha önce de belirtildiği gibi, işlemin etkisini ortaya çıkarmaktadır. Ancak bireysel farklılıklar gruplar içinde denekler arasında ve gruplar arasında farklılıklar yaratır ki her ikisi de işleme ilgili değildir. Bu bireysel farklılıklar ortak değişken olarak ölçülür ve kovaryans analiziyle ayarlanarak ortadan kaldırmaya çalışılır. Böylece deneysel kontrolün yeterli olmadığı yerlerde istatistiksel kontrol sağlanmakta ve işlemin net etkisi ortaya çıkarılmaya çalışılmaktadır.

Kovaryans analizinin klasik deneysel kullanımında ele alınan ortak değişken genellikle, işlemden önce bağımlı değişken gibi ölçülen ön test puanlarıdır. Bununla beraber bir ya da birden çok olabilen ortak değişken bağımlı değişkenden tamamen farklı bazı demografik özellikler (eğitim düzeyi, sosyo- ekonomik düzey, v.b.) veya kişilik özellikleri de (kaygı düzeyi, zeka düzeyi v.b.) olabilir.

Örneğin istatistik derslerinde sınav kaygısını azaltmak için geliştirilen yöntemleri inceleyen bir araştırma yapıldığını varsayalım. İstatistik dersini alanlar üç işlem grubuna yansız olarak atanır. Üç işlem grubu bağımsız değişkenin üç düzeyidir. Bu gruplara ortak değişken olarak standardize edilmiş kaygı testi verilir. Sonra belirli bir süre işlem uygulanır. İşlem süresinden sonra her üç gruptaki öğrencilere sınav kaygısını ölçen bir test daha verilir. (Hatırlamanın etkisinden kaçınmak için ilk testten ayrı bir formda olmalıdır). Bu ölçüm bağımlı değişken olarak hizmet verir. İstatistiksel test, ilk testteki farklılıklar ayarlandıktan sonra işlemden sonra sınav kaygısı ortalamalarının her üç grupta da aynı olduğu null hipotezidir.

2. Çoklu varyans analizinden sonra sıklıkla bağımsız değişkenler arasındaki manidar farklılıklara çeşitli bağımlı değişkenlerinin katkısının yorumlanması istenir. Bu katkıların yorumlanmasında kovaryans analizinden yararlanır.
3. Kovaryans analizinin üçüncü kullanımı, deneklerin işlemler için yansız olarak atanmadığı deneysel olmayan durumlar içindir. Bu gibi durumlarda yorumunda güçlük olmasına rağmen bir istatistiksel karşılaştırma işlemi olarak kovaryans analizi yapılır. Bu analiz temelde, eğer bütün deneklerin puanları ortak değişen ya da değişkenler de aynı yerde ise ayarlanmış grup ortalamalarının ne olduğunu bulmakta kullanılır. Ortak değişkende denekler arasındaki farklılıklar ortadan kaldırılmış olur. Sonuçta farklılığın yalnızca bağımsız değişkenin etkileri ile ilişkileri olduğu tahmin edilebilir. (Tabi farklılıklar ortak değişken olarak ele alınmayan özelliklerden dolayı da olabilir).

* M.Ü. Atatürk Eğitim Fakültesi Eğitim Bilimleri Bölümü Öğretim Üyesi

Örneğin politik tutumların (bağımlı değişken) coğrafi bölgeye (bağımsız değişken) göre değişip değişmediği araştırılmak istendiğinde politik tutum ve coğrafi bölgeyle değişmesi beklenen iki değişken sosyo-ekonomik düzey ve yaşır. Bu iki değişken ortak değişken olarak alınır. İstatistiksel analiz sosys-ekonomik düzey ve yaş ayarlandıktan sonra coğrafi bölgeye göre politik tutumların değişmediği null hipotezini test eder. Ancak eğer yaş ve sosyo-ekonomik düzey coğrafik bölgeye sıkı sıkıya bağlı ise onları ayarlamak gerçekçi değildir. Buna ek olarak bağımlı değişken, ortak değişken ilişkisi ve ortak değişken ölçümündeki güvensizlik de sonuçları yanlış yola sevk eder.

Birçok araştırmada bazı değişkenlerin kontrolü ve ele alınan değişkenin etkisinin daha net olarak ortaya çıkarılmasının önemi, daha önce de belirttiği gibi çok büyüktür. Özellikle sosyal bilimlerde yapılan araştırmalarda değişkenlerin kontrolü çok zor olmaktadır.

Eğitim bilimlerinde yapılan birçok deneysel türden araştırmada da, araştırmacı, denek gruplarının deneysel modellerin getirdiği şekilde (yansız atama) oluşturma olanağına ve hakkına sahip değildir. Buna sahip olsa bile, eğitimin yürütüldüğü ortam olan sınıflarda araştırma yapmanın önemi açıktır. Sınıflarda yapılan araştırmada araştırmacı, önceden belli (intact) grup yapısıyla bağlıdır. Popham(1967)

...Yönetici ve öğretmenlerin, öğrencileri deneysel amaçlarla kaydırmaları ve şekil vermelerinin zor olduğunu, araştırmacıya mükemmel bir model sağlamak için öğrencilerin bir öğretmenden diğerine kaydırılmalarının uygun olmadığını ve dolayısıyla araştırmacının önceden belli gruplarla çalışmaya razı olması gerektiğini...

söylemektedir. Bu durumda araştırmacının ne yapması gerektiği konusunda da Popham kovaryans denilen bir istatistiki tekniğin uygulanabileceğini, eğitim istatistikçileri için kullanışlı istatistiki olanak sağlayacağını belirtmektedir. Ayrıca kovaryans analizinin doğal olarak her akla gelen yerde kullanılmaması gerektiğini, ancak okullarda yürütülen araştırmalarda da yeteri sıklıkta kullanılmamasının üzücü olduğunu söylemektedir.

Türkiye'de kovaryans analizi eğitim bilimlerinde deneysel araştırmalarda az kullanılmakta ve daha çok deneysel araştırmalara özgü bir teknik olarak bilinmektedir. Ancak bu analiz daha önce de belirtildiği gibi deneysel olmayan durumlarda da kullanılabilir. Bu yazıda kovaryans analizi bu yönüyle ele alınmış ve örneklendirilmeye çalışılmıştır.

Kovaryans analizi bir istatistiksel kontroldür. Bu istatistiksel kontrol bazı varsayımlar yapılmasını gerektirmektedir.

VARSAYIMLAR

Kovaryans analizi varyans analizi ile regresyon analizinin birlikte kullanılması olduğundan, gerek varyans analizinin, gerekse regresyon analizinin varsayımlarının yapılmasını zorunlu kılar. Bunun için bu varsayımlar varyans analizinin ve regresyon analizinin varsayımları olarak ele alınmıştır.

1. Varyans analizinin varsayımları (Winer, 1971; Ferguson, 1971; Kirk, 1968; Howell, 1992):

a) Normallik varsayımı: her grup içinde bağımlı değişken normal dağılmıştır veya en azından fazla çarpık değildir. Genelde varyans ve kovaryans analizlerinin normal dağılımdan sapmalara karşı dayamlı olduğu literatürlerde belirtilmektedir.

b) Hata varyanslarının homojenliği: Grupların hata varyansları birbirine yakın olmalıdır. Bu varsayımdan küçük sapmaların önemli olmadığı düşünülmektedir. Ayrıca sapmaların önemiyle ilgili testler de geliştirilmiştir. Bunlar; iki grup karşılaştırıldığında hesaplanan F değeridir. Bu F değeri büyük olan varyansın küçük olan varyansa bölünmesi ile elde edilir. İki'den daha çok gruba karşılaştırılmalarında da grupların eleman sayılarının farklı olup olmadığına göre hesaplanan Bartlett testidir. Varyansların homojenliği konusunda bir şüphe varsa bu testlerden uygun olanla test edilmelidir. Homojen bulunmuşsa verilerin transformasyonu yoluna gidilmelidir.

c) Hataların bağımsızlığı : Bağımlı değişkende sistematik hata kaynakları olmamalıdır. Varyans analizinde, gruplar arasında oluşan sistematik hatalar kovaryans analizi ile düzeltilmelidir denilmektedir.

Kovaryans analizi yapıldığından bu hatalar bu testle düzeltilmektedir. Kovaryans analizi için bu varsayım gerekmemektedir.

2. Regresyon analizinin dayandığı varsayımlar (Kirk, 1968, Tabachnick Fidell, 1989):

a) Regresyonun homojenliği: Her grup için ayrı ayrı elde edilmiş regresyon katsayılarının birbirine yakın olması gerekmektedir. Bağımsız değişken(ler) ile ortak değişken(ler) arasında bir ortak etki sözkonusu ise bu regresyonun heterojen olduğunu göstermektedir.

b) Ortak değişken(ler) ile bağımlı değişken arasındaki ilişki doğrusal veya doğrusala yakın olmalıdır. Bu varsayım da zihinsel ve eğitsel testlerde ve özellikle denek sayısının fazla olması durumunda genellikle karşılanabilmektedir.

Kovaryans analizi ile ilgili sonuçlar bu varsayımların yerindeligi oramnda geçerli olacaktır.

Kovaryans Analizinin Sınıflandırılması ve Kullanılan Değişkenler

Kovaryans analizi tekli ve çoklu olarak sınıflandırılmaktadır. Tekli kovaryans analizinde bir bağımlı, bir bağımsız (iki veya daha çok grup olabilir) ve en az bir kontrol değişkeni vardır. Çoklu kovaryans analizinde bir bağımlı, birden çok bağımsız değişken (herbiri alt gruplarla temsil edilen) ve bir ya da daha çok kontrol değişkeni vardır. Bağımlı değişken (Y), kontrol değişkeni ((X)leri (X_1, X_2, X_3, vb) sembolü ile gösterilir. Bağımsız değişken gruplarla temsil edilir. Genelde değişkenlerin bağımlı, bağımsız, kontrol şeklinde tanımlanmasına deneysel araştırmalarda rastlanılmaktadır. Burada deneysel olmayan araştırmada da aynı şekilde tanımlanmıştır.

Kontrol değişkenlerinin seçimi

Kontrol değişkeninin seçiminde bu değişkenin bağımlı değişkenle ilişkili olması(yansız atama yapılan araştırmalarda $r \geq 0.30$, yansız atama yapılan araştırmalarda $r > 0.30$), eğer birden fazla kontrol değişkeni varsa bunların birbiri ile ilişkili olmaması istenir(Frigon, Laurencelle, 1993). Bir kontrol değişkeni alınacaksa bağımlı değişkenle en yüksek ilişkiye sahip olan seçilir. Birden fazla olduğunda çoklu korelasyonun karesine bakmak yararlı olabilir. Ayrıca kontrol değişkeni bağımsız değişkene bağlı ise analiz sonuçları geçerliliğini yitirir. (Tabachnick, Fidell, 1989; Kirk, 1968).

Kovaryans Analizinin Hesaplama İşlemleri

Burada yalnızca iki grubun ele alındığı ve tek değişkenin kontrol edildiği tekli kovaryans analizi örneklendirilmiş ve hesaplama süreci Tablo 1'de verilen yapay veriler üzerinde açıklanmıştır. Tabiidir ki bu örnekteki veri sayısı oldukça küçüktür. Veri sayısının kovaryans analizinin hesaplama koşullarını karşılayabilecek kadar olması gerekir.

Örnek: 20 kişilik bir örneklemin lise not ortalamalarını (LNO) kontrol ederek klasik fen programı fen kolundan mezun öğrencilerle modern fen programı dil-edebiyat kolundan mezun öğrencilerin üniversitedeki akademik başarıları arasındaki farkı test ediniz.

Y: Bağımlı değişken olan üniversitedeki akademik başarı ortalaması

X: Kontrol edilen değişken olan LNO

Tablo 1

	G	R	U	P	L	A	R
	Fen Kolu			Dil-Edebiyat Kolu			
	Y	X		Y	X		
	47	5		75	7		
	63	7		80	9		
	37	6		72	5		
	51	6		63	8		
	29	5		34	5		
	30	5		67	8		
	38	6		52	7		
	42	7		69	7		
	57	6					
	62	7					
	53	5					
	46	7					
n_j		12			8		
$T_{yj}; T_{xj}$	555	72		512	56		
$\bar{Y}; \bar{X}$	46.25	6.00		64.00	7.00		
$\sum_{i=1}^{n_j} Y_{ij}; \sum_{i=1}^{n_j} X_{ij}$	27115	440		34288	406		
T_{xjy_j}		3384			3662		
$\frac{T_{yj}^2}{n}; \frac{T_{xj}^2}{n}$	25668.75	432.00		32768.00	392.00		

$N:20 (12+8)$

$T_y:1067 (555+512)$

$\bar{Y}:53.35 (1067/20)$

$A_y = \frac{T_y^2}{N} : 56924.45 (1067^2/20)$

$B_y = \sum_{i=1}^{n_j} \sum_{j=1}^k Y_{ij}^2 : 61403.00 (27115+34288)$

$C_y = \sum_{j=1}^k T_{yj}^2 / n_j : 58436.75 (555^2/12+512^2/8)$

$T_x:128 (72+56)$

$\bar{X}:6.40 (128/20)$

$A_x = T_x^2 / N : 819.20 (128^2/20)$

$B_x = \sum_{i=1}^{n_j} \sum_{j=1}^k X_{ij}^2 : 846.00 (440+406)$

$C_x = \sum_{j=1}^k T_{xj}^2 / n_j : 824.00 (72^2 / 12 + 56^2/8)$

Öneri, C:1 S:1

$$A = T_{xy} : 7046 (3384 + 3662)$$

$$B = \sum_{j=1}^k \frac{T_{xj} T_{yj}}{n_j} : 6914.00 ((555 \times 72) / 12) (512 \times 56) / 8))$$

$$C = \frac{T_x T_y}{N} : 6828.80 ((1067 \times 128) / 20)$$

Bu hesaplamalardan sonra gruplararası, gruplarıçi ve toplam kareler toplamları ve çarpımlar toplamları bulunur.

Kareler toplamı

	Y	X
Gruplararası Ay)	(Cy-58436.75-56924.45=1512.30 (Byy)	(Cx-Ax)824.00-819.20=4.80 (Bxx)
Gruplarıçi Cy)	(By-61403.00-58436.75=2966.25 (Sxx)	(Bx-Cx)846.00-824.00=22.00 (Sxx)
Toplam Ay)	(By-61403.00-56924.45=4478.55 (Tyy)	(Bx-Ax)846.00-819.20=26.80 (Txx)

Çarpımlar toplamı

Gruplararası (B-C)	6914.00-6828.80=85.20 (Bxy)	
Gruplarıçi (A-B)	7046.00-6914.00=132.00 (Sxy)	
Toplam (A-C)	7046.00-6828.80=217.20 (Txy)	

Hesaplanan kareler toplamı toplamsaldır. Diğer bir deyişle gruplararası ve gruplarıçi kareler toplamları toplam kareler toplamını verir. Aynı şekilde gruplararası ve gruplarıçi çarpımlar toplamı, toplam çarpımlar toplamını vermektedir.

Bu şekilde kareler toplamları ve çarpımlar toplamları bulunduğundan sonra ayarlanmış kareler toplamlarının hesaplanmasına geçilir.

Ayarlanmış Toplam Kareler Toplamı (T_{ay})

$$T_{ay} = T_{yy} - \frac{T_{xy}^2}{T_{xx}}$$

Bu formüldeki terimler daha önce hesaplanmış olan toplam kareler toplamları (Y ve X için) ve toplam çarpımlar toplamıdır.

$$T_{ay} = 4478.55 - \frac{217.20^2}{26.80} = 2718.26$$

aynı şekilde:

Ayarlanmış Gruplarıçi Kareler Toplamı (S_{ay})

$$S_{ay} = S_{yy} - \frac{S_{xy}^2}{S_{xx}} \quad S_{ay} = 2966.25 - \frac{132.00^2}{22.00} = 2174.25 \quad \text{elde edilir.}$$

Ayarlanmış Gruplararası Kareler Toplamı (B_{ay})

Ayarlanmış toplam kareler toplamından gruplarıçi kareler toplamı çıkarılarak bulunur.

$$B_{ay} = T_{ay} - S_{ay}$$

$$B_{ay} = 2718.26 - 2174.25 = 544.01$$

Bu ayarlanmış kareler toplamları F oranını hesaplamak için serbestlik derecelerini ve kareler ortalamasını içeren Tablo 2'de verilen kovaryans analizi tablosuna geçirilir.

Tablo-2

Kovaryans Analizi Tablosu

Kaynak	Ay.KT (X)	Ay. KT için ser.Ay.KO der.	F
Gruplararası	544.01	k-1:1	544.01 (Ay.KO _B) 4.25
Gruplarıçi	2174.25	N-k-1:17	127.90 (Ay.KO _S)
Toplam	2718.26	N-k:18	

Tablodaki ayarlanmış kareler ortalamaları (Ay.KO) ayarlanmış kareler toplamlarının kendi serbestlik derecelerine bölünmesiyle elde edilmiştir.

Tablodaki F değeri ise;

$$F = \frac{Ay.KO_B}{Ay.KO_S}$$

ayarlanmış gruplararası kareler ortalamasının gruplarıçi kareler ortalamasına bölünmesiyle elde edilir. Hesaplanan F oranı gruplararası ve gruplarıçi serbestlik derecelerine göre ve istenilen manidarlık düzeyinde F değerlerini veren Tablo ile karşılaştırılır. Bu örnekte 1 ve 17 serbestlik derecelerine göre .05 düzeyindeki Tablo F değeri 4.45'dir. Hesaplanan F oranı ise; 4.25'tir. Hesaplanan F oranı 4.45'ten küçük olduğu için manidar değildir.

Bir sonraki aşamada iki grubun LNO'na göre ayarlanmış akademik başarı ortalamaları aşağıdaki eşitlik yardımıyla bulunur.

$$\bar{Y}_j: bw (\bar{X} - \bar{X}_j) + \bar{Y}_j$$

$$\text{Eşitlikte } bw = \frac{S_{xy}}{S_{xx}}$$

$$\text{Örnekte; } bw = \frac{132.00}{22.00} = 6.00 \text{ 'dir.}$$

Buradan, birinci grubun (fen kolu) ayarlanmış ortalaması

$$\bar{Y}_1 = 6.00 (6.40 - 6.00) + 46.25$$

$$\bar{Y}_1 : 48.65$$

İkinci grubun (dil-edebiyat) ayarlanmış ortalaması;

$$\bar{Y}_2 = 6.00 (6.40 - 7.00) + 64.00$$

$$\bar{Y}_2 : 60.40$$

olarak hesaplanır.

Bu eşitlikte \bar{X} : Her iki grup birlikte ele alındığında LNO

\bar{X}_j : Her bir grubun LNO

\bar{Y}_j : Her bir grubun akademik başarı ortalamasıdır.

İki grubun ayarlanmış ortalamaları karşılaştırılacak olursa; Fen kolunda 46.25 iken 48.65; dil-edebiyat kolunda ise 64.00 iken 60.40 olmuştur. Analiz sonunda F oranı manidar bulunmamış, ayrıca LNO daha düşük olan grup olan fen kolunun akademik başarı ortalaması ayarlandıktan sonra daha yüksek, diğeri ise daha düşük bulunmuştur. Buradan akademik başarıdaki baştaki farklılığın LNO'nun etkisine bağlı olduğu sonucu çıkartılabilir. Bu etki F oranına da yansımıştır. Aynı veriler için akademik başarı ortalamaları arasındaki farkın manidarlığını test etmek için yapılan tek boyutlu varyans analizi

sonucunda F oranı (t^2) 9.18 olarak bulunmuştur. Bu F oranı 1 ve 18 serbestlik derecelerine göre .01 düzeyinde (8.28) manidardır.

Kovaryans analizinde ikiden daha çok grup karşılaştırıldığında F oranı manidar bulunursa hangi gruplar arasında manidar fark olduğunu bulmak gerekir. Bunun için çeşitli testler vardır. Bunlar; en küçük önemli farklar, Tukey gerçek önemli farklar, Scheffe'nin t testi Newmann Keuls test ve Duncan'ın çoklu genişlik testidir(Kirk, 1968).

KAYNAKLAR

- Ferguson, G.A. Statistical Analysis in Psychology and Education. New York: McGraw-Hill, 1971.
- Frigon, J_Y and L. Laurencelle., "Analysis of Covariance: A Proposal Algorithm." Educational and Psychological Measurement, 1993 53:1-18
- Goehring, Harvey, J. Jr. Statistical Methods in Education, Virginia : Information Resources Press, 1981
- Guenther, W.C. Analysis of Variance, New Jersey: Prentice-Hall, 1964
- Howell, C.David. Statistical Methods for Psychology. Third Edition, California : Duxbury Press, 1992.
- Kirk, E. Roger; Experimental Design: Procedures for the behavioral sciences. California: Cole Publishing Company, 1968
- Popham, W. James. Educational Statistics, Use and Interpretation , Los Angeles: Harper And Row Publishers, 1967
- Tabacnick, G.B., Abd L.S.Fidell, Using Multivariate Statistics, Second edition, Northridge:Harper Collins Publishers, 1969
- Winer, B.J. Statistical Principles in Experimental Design, Second Edition, New York; McGraw-Hill Book Company, 1971