

SORGULAYICI ve ONAYLAYICI FAKTÖR ANALİZİNİN KARŞILAŞTIRILMASI

Dr. Mürüvvet PAMUK*

ÖZET

Bu çalışmada, sorgulayıcı (exploratory) ve onaylayıcı (confirmatory) faktör analizi tekniklerinin bir karşılaştırılması verilmiştir. Önce, sorgulayıcı ve onaylayıcı faktör analizi metodlarının anahatları açıklanmış ve bu metodların karşılaştırılması yapılarak, açıklayıcı bir örnek verilmiştir.

ABSTRACT

This paper presents a comparison between exploratory and confirmatory factor analysis techniques. The basic processes of constructing confirmatory and exploratory factor analysis methods are described. Then, the comparison of exploratory and confirmatory factor analysis is discussed and an illustrative example is given.

* Okutman, İ.Ü. İktisat Fakültesi

1) GİRİŞ

Faktör analizi yöntemi, ilk defa XX.yüzyılın başında Spearman (1904) tarafından geliştirilmiştir (Stapleton, 1997 (1), s.1). Önceleri yalnızca boyut indirgeme amacıyla kullanılırken, bilgisayarların gelişmesiyle faktör analizinde de ilerlemeler kaydedilmiş ve kurulan hipotezlerinde testi yapılabilmektedir. Faktör analizi yaklaşımı, yapılan çalışmanın özelliğine bağlı olarak farklı şekillerde ele alınabilir.

Bu çalışmada, faktör analizine farklı yaklaşım şekilleri üzerinde durulmuş ve sayısal bir örnek üzerinde yaklaşımlar açıklanmıştır.

2) SORGULAYICI FAKTÖR ANALİZİ

Rennie (1997, s.1) faktör analizinin iki boyutta ele alınabileceğini belirtmiştir: Bunlar "teori geliştirme" ve "teori değerlendirme"ye yöneliktir. Hangisinin kullanılacağı yapılacak olan veri analizinin amacına göre belirlenir. Sorgulayıcı faktör analizi, çok sayıda değişkenleri temsil eden daha az sayıda hipotetik faktörleri belirlemek amacıyla kullanılır. Analize bu faktörlerin neler olduğu hakkında ön bilgiler olmaksızın değişkenler arasındaki korelasyonların belirlenmesi ile başlanır (Stapleton, 1997 (2), s.6).

Sorgulayıcı faktör analizi, değişkenleri kapsayan faktör sayılarının belirlenmesine ve faktörlerin ilişkili mi yoksa ilişkisiz mi olduğunu belirlemeye yardımcı olur.

Sorgulayıcı faktör analizinin çeşitli dezavantajları vardır. Stapleton makalesinde Mulaik'in "belli ön varsayımlarda bulunmaksızın deneylerden bilgi elde etmenin rasyonel olmadığını" söylediğini belirtir (Stapleton, 1997 (1), s.2). Sorgulamaya yönelik varsayımlar verilen veri kümesindeki değişkenler arasındaki ilişkileri daima belirtmeyebilir de. Genel faktör analizi modeli bir doğrusal modeldir ve yalnız belli veri çeşitleri için uygundur. Halbuki, birçok nedensel ilişkiler doğrusal değildir. Stapleton "doğrusal bir ilişkinin sonuçlar üretmesinin harikulade olabileceğini, fakat bu sonuçların aldatıcı olabileceğini" belirtir (1997 (1), s.2).

Sorgulayıcı faktör analizi ile elde edilen faktör yapıları özel teorilere ve döndürme tekniklerine bağlıdır. Bu da, hassas sonuçlar üretebilir.

Bir başka sorun, sorunların yorumlanmasındadır. Birkaç değişkenle ölçülen faktörlerin yorumu zor olabilir. Stapleton (1997 (1), s.2) makalesinde

Mulaik'in "yorumlamadaki zorluğun araştırmacının ön bilgiye sahip olmamasına ve bu nedenle yorum yapmak için bir temel oluşturulmamasına bağlı olduğunu" belirttiğini yazar.

3) ONAYLAYICI FAKTÖR ANALİZİ

Sorgulayıcı faktör analizi gibi teori geliştirmeye karşılık, onaylayıcı faktör analizi bir teori test etme yöntemidir. Bu yöntemde araştırmacı analizine ön bir hipotezle başlar. Bu model veya hipotez, hangi faktörlerin ilişkili olduğunu ve hangi değişkenlerin hangi faktörlerle ilişkili olduğunu belirtir. "Hipotez kuvvetli bir teorik ve/veya ampirik temel üzerine kurulur" (Stapleton, 1997 (1), s.3).

Ön faktörler belirlendikten sonra, onaylayıcı yöntemler, önceden belirlenen faktör modelinin uygunluğunu belirlemek için gözlenen ve teorik faktör yapılarının birbirine uyup uymadığını belirlemeye çalışır.

Faktör katsayıları, faktör korelasyon katsayıları ve ölçüm hatalarının varyans/kovaryansı gibi sabit ve serbest özel parametrelerle farklı modeller belirlenir. Bu parametreler araştırmacının teorik beklentisine göre kurulur. Gillapsy sabit ve özel parametreleri şöyle tanımlar (Stapleton, 1997 (1), s.3):

"Bir parametreyi sabit tutmak, parametreyi kişinin beklentisine bağlı olarak belli bir değere ayarlamak demektir. Böylece, bir parametreyi sabit tutmakla araştırmacı parametrenin analizde tahmin edilmesine izin vermez. Bir parametreyi serbest tutmak ise, modeli veriye uydurmak için analiz sırasında parametrenin tahmin edilmesine izin verilmesi demektir". Daha sonra modeller bilgisayar programı yardımıyla bir diğerine göre test edilir. Analiz tamamlandığında, modellerin veriye ne derece uyduğunu belirlemek veya değişkenler arasındaki kovaryansı açıklamak için çeşitli farklı istatistikler elde edilir. Elde edilen istatistikler "uygunluk istatistikleri" olarak tanımlanır. Uygunluk istatistikleri parametrelerin tümünü test eder. Bu uygunluk istatistikleri, gözlenen ve gizli değişkenler arasındaki ilişkileri, önceden belirlenen modellerin hangisinin en iyi açıkladığını belirlemek için kullanılır.

Asıl problem, ölçülen değişkenler üzerinde modelin örnek veriye uygunluğunu belirlemek ve modelin parametrelerini en iyi şekilde tahmin etmektir. Eğer model veriye uymazsa, önerilen model gözlenen değişkenlerle ilgili nedensel yapının belirlenmesinde olası bir seçim olarak reddedilir. Eğer model istatistiksel olarak reddedilmezse, nedensel yapının uygun bir sunumu olarak kabul edilir.

4) ONAYLAYICI FAKTÖR ANALİZİNE KARŞI SORGULAYICI FAKTÖR ANALİZİ

Her iki faktör analizindeki temel matematiksel hesaplamalar birbiriyle ilgilidir. Sorgulayıcı faktör analizi, değişkenler kümesine ait modeli veya faktör yapısını belirlemeyi amaçlar. Bu işlem de faktör sayısını ve faktör yüklerinin örüntüsünü belirlemeyi kapsar. Sorgulayıcı faktör analizi genellikle teori üretme yöntemi olarak düşünülür. Aksine, onaylayıcı faktör analizi genellikle kuvvetli bir teorik ve/veya ampirik bulgular üzerine kurulur. Bu bulgular da, araştırmacının ileri aşamada faktörü tam belirlemesini kolaylaştırır. Bu faktör modeli, hangi faktörlerin ilişkili olduğu kadar hangi değişkenlerin hangi faktörler üzerinde yüklemde bulunacağını belirler. Onaylayıcı faktör analizinde, faktörlerle ilgili çeşitli hipotezleri test etmek mümkün olduğundan, sorgulayıcı faktör analizine göre avantajlıdır.

5) AÇIKLAYICI ÖRNEK

Sorgulayıcı ve onaylayıcı faktör analizinin çeşitli özelliklerini açıklamak için basit bir sayısal örnek ele alınmıştır. Kullanılan veri matrisi, Devlet İstatistik Enstitüsü (DİE) tarafından yayınlanan Ocak 1995- Temmuz 2000 yılları arasında 1992=100 alınarak hesaplanan aylık imalat sanayi üretim endeksi değerleridir. İmalat Sanayi değişkenleri x_1 =gıda,içki,tütün imalatı, x_2 =tekstil, giyim ve deri sanayi, x_3 =kağıt,kağıt ürünleri ve basım, x_4 =kimya sanayi, x_5 =taş-toprağa dayalı imalat, x_6 =metal ana sanayi, x_7 =metal eşya,makine ve teçhizat imalatı, x_8 =elektrik,gaz ve su imalatı'dır. Faktör analizinin ilk aşaması olarak değişkenlere ait korelasyon matrisi Tablo:1'de verilmiştir.

Tablo:1 Sekiz İmalat Değişkeni Arasındaki Korelasyonlar

Değişkenler (j)	1	2	3	4	5	6	7	8
1- Gıda, içki, tütün	1,000	0,164	0,128	0,217	0,196	0,235	0,293	0,111
2- Tekstil, giyim ve deri	0,164	1,000	0,358	0,119	-0,132	-0,084	-0,262	-0,568
3- Kağıt, kağıt ürün ve basım	0,128	0,358	1,000	0,094	0,134	0,063	-0,005	-0,216
4- Kimya	0,217	0,119	0,094	1,000	0,424	-0,234	-0,200	-0,190
5- Taş-toprağa dayalı imalat	0,196	0,132	0,134	0,424	1,000	0,295	0,368	0,058
6- Metal ana sanayi	0,235	-0,084	0,063	-0,234	0,295	1,000	0,614	0,577
7- Metal eşya, makine ve teç	0,293	-0,262	-0,005	-0,200	0,368	0,614	1,000	0,707
8- Elektrik, gaz ve su	0,111	-0,568	-0,216	-0,190	0,058	0,577	0,707	1,000

Analizin ikinci aşaması, temel bileşenler analizi kullanılarak korelasyon matrisini faktör matrisine çevirmektir.

Tablo: 2'den görüldüğü gibi, korelasyon matrisinin 1'den büyük özdeğerlerinin sayısı 3'dir. Bu nedenle, temel bileşenler analizi üç bileşen üretmektedir.

Tablo:2 Açıklanan Toplam Varyans

Bileşenler	Özdeğerler	Varyans %	Birikimli %
1	2,662	33,3	33,3
2	1,879	23,5	56,8
3	1,161	14,5	71,3
4	0,815	10,2	81,5
5	0,666	8,3	89,8
6	0,391	4,9	94,7
7	0,293	3,7	98,4
8	0,133	1,6	100,0

Tablo: 3'deki temel bileşen yükleri, değişkenlerin temel bileşenle olan korelasyonlarıdır. Bu yüklerin kareleri temel bileşenin o değişkenin varyansını ne ölçüde açıkladığını gösterir. Birinci temel bileşene ait kareli yüklerin toplamı birinci özdeğere (λ_1) eşittir. Standardize edilmiş sekiz değişkenin toplam varyansı 8'dir. $\lambda_1=2,662$ olduğundan, bulunan birinci temel bileşen toplam varyansın %33,272'sini açıklar. İkinci ve üçüncü temel bileşenler de sırasıyla toplam varyansın %23,490 ve %14,512'sini açıklarlar. Modelin toplam varyansın %71,274'ünü açıkladığı söylenebilir.

Temel bileşenler yöntemiyle elde edilen bu ilk çözüme dik döndürme yöntemlerinden Varimax yöntemi uygulandığında, Tablo:4'de verilen faktör matrisi elde edilir. Görüldüğü gibi faktör yapısı daha basit bir yapıya indirgenmiştir. Dik döndürmeden sonra ise, faktörler sırasıyla toplam değişimin %31,607, %20,122 ve %19,545'ini açıklamaktadır. Dik döndürme sonucu, birinci faktör son üç değişkeni, ikinci faktör ikinci ve üçüncü değişkenleri, üçüncü faktör ise birinci, dördüncü ve beşinci değişkenleri içermektedir.

Tablo:3 Temel Bileşenler Matrisi

Değişkenler	Bileşenler			h ²
	1	2	3	
1- Gıda, içki, tütün	0,263	0,576	5,771E-03	0,401
2- Tekstil, giyim ve deri	-0,490	0,563	0,430	0,742
3- Kağıt, kağıt ürün. ve basım	-0,167	0,544	0,519	0,593
4- Kimya sanayi	-0,258	0,557	-0,694	0,859
5- Taş- toprağa dayalı imalat	0,278	0,706	-0,345	0,694
6- Metal ana sanayi	0,778	0,226	0,300	0,746
7- Metal eşya, makine ve teçh.	0,884	0,186	7,668E-02	0,822
8- Elektrik, gaz ve su	0,892	-0,200	-9,965E-02	0,845
Özdeğerler (λ)	2,662	1,879	1,161	5,702
Toplam varyans (%)	33,272	23,490	14,512	71,274

Faktörlerin içerdiği bu değişkenler Tablo: 1'deki korelasyon matrisinde de görüldüğü gibi birbirleriyle ilişkilidirler. Yalnız üçüncü faktörde yer alan birinci değişkenin dördüncü ve beşinci değişkenlerle ilişkisi azdır, zaten Tablo: 4'deki faktör matrisinde de birinci değişkenin yükü diğerlerine göre azdır, dolayısıyla üçüncü faktör birinci değişkenin varyansının daha az bir kısmını açıklar.

Tablo: 4 Döndürülmüş Faktör Matrisi

Değişkenler	Faktörler		
	1	2	3
1- Gıda, içki, tütün	0,352	0,291	0,439
2- Tekstil, giyim ve deri	-0,249	0,818	0,108
3- Kağıt, kağıt ürün. ve basım	7,767E-02	0,764	5,227E-02
4- Kimya sanayi	-0,333	-4,314E-02	0,864
5- Taş- toprağa dayalı imalat	0,249	0,126	0,769
6- Metal ana sanayi	0,857	0,108	8,636E-03
7- Metal eşya, makine ve teçh.	0,890	-0,107	0,133
8- Elektrik, gaz ve su	0,783	-0,481	-3679E-02
Özdeğerler (λ)	2,529	1,610	1,564
Toplam varyans (%)	31,607	20,122	19,545

Yukarıda anlatılan metodlar sorgulayıcı faktör analizi kapsamında ele alınmıştır.

Onaylayıcı faktör analizi kapsamında ise, modelin yapısı hakkında bilgi edinmek için çeşitli testler uygulanabilir. Örneğin, ele alınan veri matrisinin faktör analizi için uygun olup olmadığı bulunabilir. Faktör analizi birbiriyle ilişkili değişkenleri biraraya toplayarak daha az sayıda faktöre indirgemeyi amaçlar. Bu bakımdan, faktör analizinin uygulanabilmesi için gereklilik kısıtı değişkenlerin bağımsız olmamasıdır. Bu gerekliliği kontrol etmek için, korelasyon matrisinin birim matris olup olmadığını test etmek gerekir. Bunun için Barlett tarafından geliştirilen küresellik testi önerilir (Tatlidil, 1992, s.127). Küresellik testi için

$$H_0 : R=I$$

$$H_A : R \neq I$$

hipotezleri kurulmaktadır.

$$u = |R|, \quad v = n - 1 \quad \text{ve}$$

$$u' = -\left[v - \frac{1}{6}(2p+5)\right] \ln u \sim \chi^2_{\frac{1}{2}p(p-1)}$$

alınarak, eğer $u' > \chi^2_{\alpha, \frac{1}{2}p(p-1)}$ ise H_0 hipotezi reddedilir.

" $|R|$ (korelasyon matrisinin determinanı) 0 ile 1 arasında değişir. Eğer değişkenler ilişkisiz ise $R=I$ ve $|R|=1$ olacaktır. Diğer taraftan iki veya daha fazla değişken doğrusal bağımlı ise R matrisi tam ranklı bir matris olmayacaktır ve $|R|=0$ olacaktır. Eğer değişkenler yüksek korelasyonlara sahipse $|R|$ değeri 0'a yaklaşacaktır, aksine korelasyonlar küçük ise $|R|$ değeri 1'e yaklaşacaktır" (Rencher, 1995, s.291).

Tablo:1'deki korelasyon matrisine bakılırsa, sekiz değişkenin birbiriyle ilgili olduğu görülmektedir. Bu sekiz değişkenden birbiriyle ilgili olanları birleştirerek birkaç faktöre indirgemek yararlı bir boyut indirgeme olabilir. Bunu test etmek için

$$u = |R| = 0,0479, \quad v = 67-1=66 \quad \text{ve} \quad p=8 \quad \text{alınarak}$$

$$u' = -\left[66 - \frac{1}{6}(16+5)\right] \ln(0,0479)$$

$$= -(62,5)(-3,037)$$

$$= 189,847$$

elde edilir. χ^2 tablosundan $\chi^2_{0,05;28} = 41,337$ bulunur. $189,847 > 41,337$ olduğundan, korelasyon matrisinin birim matris olduğunu belirten H_0 hipotezi reddedilir. Böylece bu korelasyon matrisinin faktör analizi için uygulanabilir olduğu sonucuna varılır.

Bazı yazarlar, faktör analizi modelini başarılı bir şekilde uygulayabilmek için R^{-1} (orijinal korelasyon matrisinin tersi)'nin köşegen matrise yakın olması gerektiğini belirtirler. R^{-1} 'in köşegen matrise nasıl yaklaştığını belirlemek için, "Kaiser tarafından örnek uygunluğu için başparmak kuralı endeksi önerilmiştir" (Rencher, 1995, s.483). Bu endeks 'örnek uygunluğu ölçümü'

(measure of sampling adequacy) olarak isimlendirilir ve bu amaç için kaba bir fikir verir.

$$MSA = \frac{\sum_{i \neq j} r_{ij}^2}{\sum_{i \neq j} r_{ij}^2 + \sum_{i \neq j} q_{ij}^2}$$

Burada, r_{ij}^2 , R 'nin i . ve j . elemanının karesi ve q_{ij}^2 ise $Q=DR^{-1}D$ matrisinin ($D=[(\text{diag}R^{-1})^{1/2}]^{-1}$ olmak üzere) i . ve j . elemanının karesidir. Bu endeks 0 ile 1 arasında değişir. Gerçektende R^{-1} matrisinin tüm köşegen dışı elemanları sıfır ise bu endeks değeri 1 olacaktır. Tablo:1'de verilen korelasyon matrisi için MSA endeksi hesaplanmış ve 0,52 bulunmuştur. Bu değer kötü bir sonuçtur (Kim and Mueller, 1988, s.54). Barlett küresellik testi sonucuna göre korelasyon matrisinin faktör analizi için uygulanabilir olduğunu belirtmiştik. MSA endeksine göre ise kötü denebilecek bir sonuç elde edilmiştir. Daha iyi bir sonuç elde edebilmek için veri matrisinin boyutunu ve korelasyon değerlerinin ortalama büyüklüklerini artırmak gerekir.

6) SONUÇ

Yapılan çalışmanın özelliğine göre, faktör analizi kapsamında sorgulayıcı veya onaylayıcı veya her iki teknikte ele alınabilir. Verilen örnekte de görüldüğü gibi, veri matrisinin çözümlenmesi olarak, önce değişkenlere ait korelasyon matrisinin bulunması, temel bileşenlerin elde edilmesi, bunlara faktör döndürmesi uygulanarak daha anlamlı faktörler (bileşenler) elde edilmesi sorgulayıcı faktör analizini kapsamaktadır. Eğer modelin yapısı ile ilgili bilgiler elde edilmesi isteniyorsa, o zaman onaylayıcı faktör analizi ile ilgili çeşitli testlerin uygulanması gereklidir.

Kaynakça

- Kim, J.O., and Mueller, C.W. (1988). **Factor Analysis: Statistical Methods and Practical Issues**. Beverly Hills, Sage Publications.
- Rencher, A.C. (1995). **Methods of Multivariate Analysis**. John Wiley & Sons, Inc.
- Rennie, K.M. (1997, January). **Exploratory and Confirmatory Rotation Strategies in Exploratory Factor Analysis**. <http://ericae.net/ft/tamu/Rota.htm>. (23/07/2000).
- Stapleton, C.D. (1997, January (1)). **Basic Concepts and Procedures of Confirmatory Factor Analysis**. <http://ericae.net/ft/tamu/Cfa.htm>. (22/07/2000).
- Stapleton, C.D. (1997, January (2)). **Basic Concepts in Exploratory Factor Analysis as a Tool to Evaluate Score Validity: A Right-Brained Approach**. <http://ericae.net/ft/tamu/Efa.htm>. (29/07/2000).
- Tatlıdil, H. (1992). **Uygulamalı Çok Değişkenli İstatistiksel Analiz**. Ankara.