

## TRAFİK KAZALARININ İNCELENMESİNDE UYUM ANALİZİNİN KULLANIMI

M. ERİŞOĞLU\* A. PEKGÖR\* A.GENÇ\*  
D. SEZER\* N. İYİT\*

### ÖZET

*Uyum analizi, satırlar ve sütunlar arasındaki uyumun ölçülmesi ile elde edilen çok boyutlu veya iki boyutlu çarpımsal ( olumsallık, çapraz sınıflandırılmış, ki-kare ) tablolarının daha az boyutlu bir düzlemde grafiksel olarak gösterimini sağlayan geometrik temelli bir istatistik tekniğidir. Uyum analizi, değişken gruplarına ait düzeyler arasındaki ilişkiyi görsel olarak sunması açısından son yıllarda kategorik verilerin analizinde sıklıkla yararlanılan tekniklerden biridir. Özellikle pazar araştırmalarında kullanımına çok rastlanan uyum analizi bu çalışmada trafik kazalarının incelenmesinde kullanılacaktır.*

*Anahtar Kelimeler : Çarpımsal Tabloları, SVD, Trafik Kazaları Uyum Analizi.*

### 1. GİRİŞ

Değişkenler arasındaki ilişkiyi inceleyen tekniklerden biri olan uyum analizinin başlangıcı Fransız Jean Paul Benzécri'nin 1960-1970'li yıllarda yaptığı çalışmalarıdır. Lebart, Morineua ve Tabard, (1977) analizini gelişim sürecinde önemli katkılarda bulunmuşlardır. Analiz ile ilgili ilk kapsamlı çalışma 1984 yılında Michael J. Greenacre tarafından yazılan "Theory and Applications of Correspondence Analysis" adlı kitap olup, bu kitapta analiz ile ilgili detaylı tanımlamalara, hesaplamalarla ilgili ayrıntılara ve uygulamalara yer verilmiştir. Fakat analiz son yıllarda popüleritesinin artmasında Carrol, Green, ve Schaffer, (1986); Hoffman ve Franke, (1986)'nın çalışmaları etkili olmuştur.

Pazar araştırmalarında, genetik araştırmalarda ve davranış bilimlerinde sıklıkla kullanılan uyum analizi bu çalışmada trafik kazalarının incelenmesinde kullanılacaktır. Çalışmada kaza sonucu türü ile kazaya sebebiyet veren kusur türleri ve kaza sonucunun türü ile kazanın olduğu yolun durumu arasında oluşturulan çarpımsal tablolara uyum analizi uygulanarak değişkenlerin düzeylerinin birbirlerine göre durumları grafiksel olarak verilecektir. Veriler, Emniyet Genel Müdürlüğü Trafik Hizmetleri Başkanlığı tarafından hazırlanan "2001 Trafik İstatistik Yıllığı"ndan alınmıştır.

### 2. UYUM ANALİZİ

Uyum Analizi (Correspondence Analysis, CA) çarpımsal tablo durumuna getirilmiş kategorik verilerin satır ve sütunlarının birlikte değişimlerini, daha az boyutlu

bir uzayda grafiksel olarak göstermeyi amaçlayan çok değişkenli analiz yöntemidir. Uyum analizi çarpımsal tablosunun ağırlıklı ana bileşenler analizi olarak da ele alınabilecek bir yöntemdir.

Veri analizine grafiksel bir yaklaşım olarak da isimlendirilen uyum analizi, iki değişken grubuna ait düzeyler arasındaki ilişkiyi görsel olarak sunması açısından çarpımsal tablolarına üstünlük sağlar.

$I \times J$  biçiminde iki boyutlu ya da  $I \times J \times K$  ve  $I \times J \times K \times L \times \dots$  biçiminde çok boyutlu olarak tablolaştırılabilen kategorik ya da kategorize edilmiş sürekli değişkenlerin kategorileri arasındaki birlikte değişimleri, tablo gözelerinin kare değerlerinden ya da değişkenlerin kategorileri arasındaki öklid uzaklıklarından yararlanılarak hesaplanan varyans ögesi (inertia) değerlerinden yararlanılarak grafiksel gösterim aracılığı ile incelemeyi amaçlayan bir tekniktir. Uyum analizi, her bir değişkenin kategorileri arasındaki ilişkileri ve değişkenler aralarındaki genel ve kategori bazında çapraz ilişkilerin grafiksel formda incelenmesini sağlar.

Uyum analizi, satır ve sütunlar arasındaki ilişkinin kolay anlaşılması için düşük boyutlu uzayda elde ettiği grafiksel gösterimin, optimal gösterim olduğunu iddia eder. Örneğin  $16 \times 16$  'lık çarpımsal tablosunun mükemmel gösterimini 15 boyutlu uzayda gösterme yerine, belki %75'lik bir bilgiye sahip olunarak iki boyutlu uzayda gösterilebilir. Bu basit gösterimi elde etmek için %25'lik bir bilgi kaybı kabul edilebilir. 15 boyutlu gösterimle sağlanan %100 'lük bilgiye karşı 2 boyutlu gösterimin sağladığı %75'lik bilgi büyük bir kazanç olabilir.

Uyum analizinde ilk aşama satır ile sütunlar arasındaki bağımlılığın istatistiksel anlamlılığının test edilmesidir. Bunun için iki yaklaşım vardır. Birinci yaklaşım, uyum analizi sonucu elde edilen yeni eksenlerin varyans ögelerinin toplamından elde edilen İz (trace) değerinin sınanmasıdır. İz'in karekökü, satır ile sütunlar arasındaki korelasyon katsayısı gibi yorumlanabilir. İz'in karekökünün 0.2 'den büyük olması ilişkinin istatistiksel olarak anlamlı olduğunu gösterir (Bendixen, 1996).

Satır ile sütunlar arasındaki ilişkinin sınanmasında ikinci yaklaşım ki-kare sınamasıdır. İz'den yola çıkarak  $\chi^2$  istatistiği aşağıdaki formül yardımı ile kolaylıkla hesaplanır.

$$\chi^2 = N \times iz \quad (1)$$

Eşitlikte yer alan  $N$  toplam sıklığı ifade etmektedir. Bilindiği gibi, elde edilen test istatistiği değeri,  $(r-1)(c-1)$  serbestlik derecesi ve  $\alpha$  anlam düzeyindeki ki-kare tablo değerinden büyük ise satırlar ile sütunlar arasındaki ilişkinin istatistiksel olarak anlamlı olduğu anlamına gelir.

Analizde ikinci aşama ise çözümlemede kullanılacak boyut sayısının tespit edilmesidir. Örneğin  $14 \times 8$ 'lik bir çarpımsal tablosunda, satırlar ve sütunların birbirinden bağımsız ve tamamen rassal olduğunu varsayarsak, 8 sütundan elde edilecek 7 boyutun herbiri olayın açıklanmasında  $(8-1)/100 = \%14.3$ 'lük bir katkı sağlar. Aynı şekilde 14 satırın gösterimini sağlayacak 13 boyutun herbiri  $(14-1)/100 = \%7.7$ 'lik katkı sağlar. Ancak uyum analizinin uygulandığı tablolarda satırlar ile sütunlar birbirleri

Bazı boyutlar daha fazla bilgi sağlarken bazıları da bilgi miktarına çok az katkıda bulunmaktadırlar. Uyum analizindeki amacın, satırlar ile sütunlar arasındaki birlikte değişimin düşük boyutlu uzayda grafiksel gösterimini elde etmek olduğu tanımlarda verilmişti. Burada boyut sayısının belirlenmesi önemli bir aşamadır. Boyut sayısına karar verilirken varyans ögelerinin iz'e bölünmesi ile elde edilen oranların birikimli toplamının %70 ve yukarısının elde edildiği boyut sayısı, uygun boyut sayısı olarak tespit edilir.

Uyum analizi, uyum tablosunun içerdiği değişken sayısına ve boyut sayısına bağlı olarak basit uyum analizi veya çoklu uyum analizi olarak iki değişik biçimde uygulanır.

## 2.1 Basit Uyum Analizi

Basit uyum analizi, I x J biçiminde gösterilen çarpımsal tablosunun ağırlıklı ana bileşenler analizini yapar. Çarpımsal tablosunun r satır ve c sütunu varsa basit uyum analizi için en büyük boyut sayısı  $d = \min((I-1), (J-1))$  dir. Uyum analizi, tablonun Pearson ki kare değerlerini ya da değişken kategorilerin birbirlerine olan öklid uzaklıklarını kullanarak elde edilen toplam varyans ögesi (total inertia) değerini parçalamayı hedefler(Özdamar, 1999).

I satır ve J sütundan oluşan, elemanları  $n_{ij}$  olan I x J boyutlu çarpımsal tablosunun i. satır toplamını  $n_{i+}$  ile j. sütun toplamını  $n_{+j}$  ile ve genel toplamı  $n_{++}$  ile gösterelim.

Çarpımsal tablosunun her gözesine ait satır ve sütun profilleri sırasıyla;

$$r_{ij} = n_{ij} / n_{i+} \quad (2)$$

$$c_{ij} = n_{ij} / n_{+j} \quad (3)$$

şeklinde hesaplanır. Satır olasılık değerleri (satır mass) ve sütun olasılık değerleri (sütun mass) ise;

$$r_i = n_{i+} / n_{++} \quad (4)$$

$$c_j = n_{+j} / n_{++} \quad (5)$$

formülleri ile belirlenir.

Tablo gözelerindeki frekansların teorik değerleri  $t_{ij}$  satır ve sütun olasılıklarından yararlanarak,

$$t_{ij} = r_i \times c_j \times n_{++} \quad (6)$$

formülü ile hesaplanır.

Elemanları  $p_{ij} = n_{ij} / n_{++}$  olan P uyum matrisinden yola çıkarak elde edilen ve elemanları  $s_{ij} = (p_{ij} - r_i c_j) / \sqrt{r_i c_j}$  olan S matrisine tekil değere dönüştürme (singular value decomposition) uygulanır (Detaylı bilgi için bkz., Press ve diğerleri, 1995).

Elemanları  $p_{ij} = n_{ij} / n_{++}$  olan P uyum matrisinden yola çıkarak elde edilen ve elemanları  $s_{ij} = (p_{ij} - r_i c_j) / \sqrt{r_i c_j}$  olan S matrisine tekil değere dönüştürme (singular value decomposition) uygulanır (*Detaylı bilgi için bkz., Press ve diğerleri, 1995*).

$$S = U\Lambda V^T \quad (7)$$

Burada,  $\Lambda$  diagonal matrisinin köşegen değerleri S matrisinin tekil değerleridir. Bu tekil değerler kullanılarak satır ve sütun koordinatları bulunur (Kurt F. ve diğerleri, 2001).

$$f_{ik} = \lambda_k u_{ik} / \sqrt{r_i} \quad k=1, 2, \dots, J \quad (8)$$

$$g_{ik} = \lambda_k v_{jk} / \sqrt{c_j} \quad k=1, 2, \dots, J \quad (9)$$

## 2.2 Çoklu Uyum Analizi

Çoklu uyum analizi veya diğer adıyla homojenlik analizi, üç veya daha fazla kategorik değişkene sahip olan çok yönlü çarpımsal tablolarının analiz edilmesinde kullanılır. Bir diğer ifade ile,  $I \times J \times K \times \dots$  şeklinde iç içe farklı şekillerde çaprazlanmış tablolarda yer alan değişkenlerin alt kategorileri arasındaki birlikteliği ve ilişkileri ortaya koymak için başvurulan grafiksel bir analizdir.

Uygulamada basit uyum analizi kullanılacağından çoklu uyum analizi ile ilgili ayrıntıya yer verilmeyecektir (Çoklu Uyum Analizi için bkz. Benzècri 1992, Gifi 1990, Nuran – Mustafa 2002, Sten- Erik Clausen 1998).

## 3. UYGULAMA

Trafik kazalarının önlenmesinde, neden sonuç ilişkisinin ortaya konması çok önemlidir. Bu doğrultuda çalışma yapılabilmesi için öncelikle verilerin doğru bir şekilde derlenmesi ve daha sonra elde edilen bu verilerin uygun tekniklerle çözümlenmesi gerekir. Bu çalışmada uyum analizi, trafik kazalarının değerlendirilmesinde kullanılacaktır. Oluşturulan iki ayrı çarpımsal tablosuna uyum analizi uygulanarak, analizin amacı olan grafiksel gösterimler elde edilecek ve bu grafikler yorumlanacaktır.

Tablo 1’de kazalara sebep olan kusurlar ve kaza türlerine ait çarpımsal tablosu verilmiştir.

**Tablo 1.** Kazalara sebep olan kusurlar ve kaza türleri

	Ölümlü Kaza	Yaralanmalı Kaza	Maddi Hasarlı Kaza	Toplam
Sürücü Asli	1089	25984	333515	360588
Sürücü Diğer	1617	36660	104531	142808
Yolcu	50	778	0	828
Yaya	635	11751	0	12386
Araç	49	820	769	1638
Yol	70	1142	469	1681
Toplam	3510	77135	439284	519929

**Tablo 2.** Tablo 1 için çarpımsal tablosu analizi

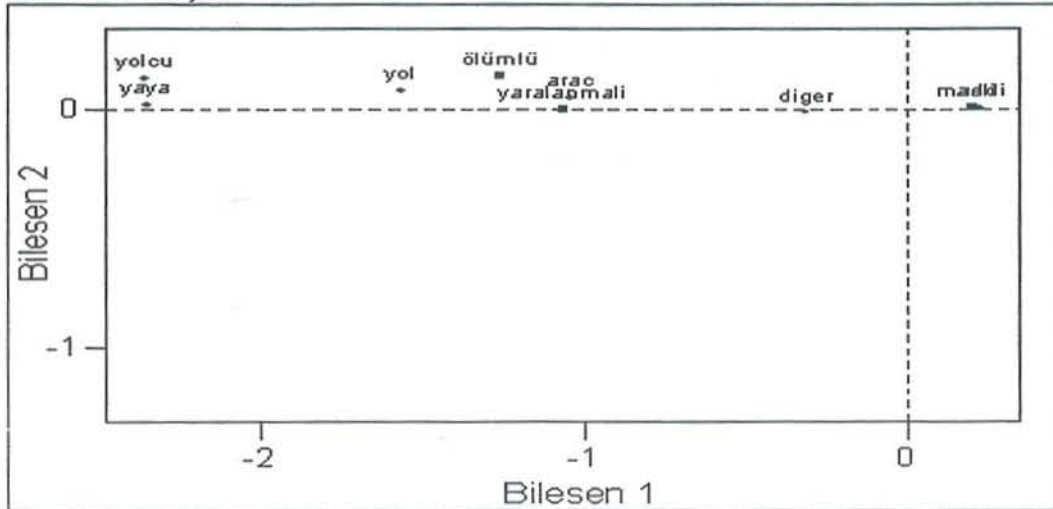
Bileşen	Varyans Ögesi	Oran	Birikimli
Bileşen 1	0,2106	0,9994	0,9994
Bileşen 2	0,0001	0,0006	1,0000
İz	0,2107	$\chi^2$	109549,04

Uyum analizi sonucunda elde edilen bileşenler değişim tamamını birlikte açıklayabilmektedirler. Bileşen 1 değişimin %99,94'ünü açıklayabilmektedir. Olayı açıklamada Bileşen 1 tek başına yeterlidir. Uyum analizinden elde edilen satır, sütun profilleri ve satır, sütun olasılıkları düzenlenerek Tablo 3'de verilmiştir.

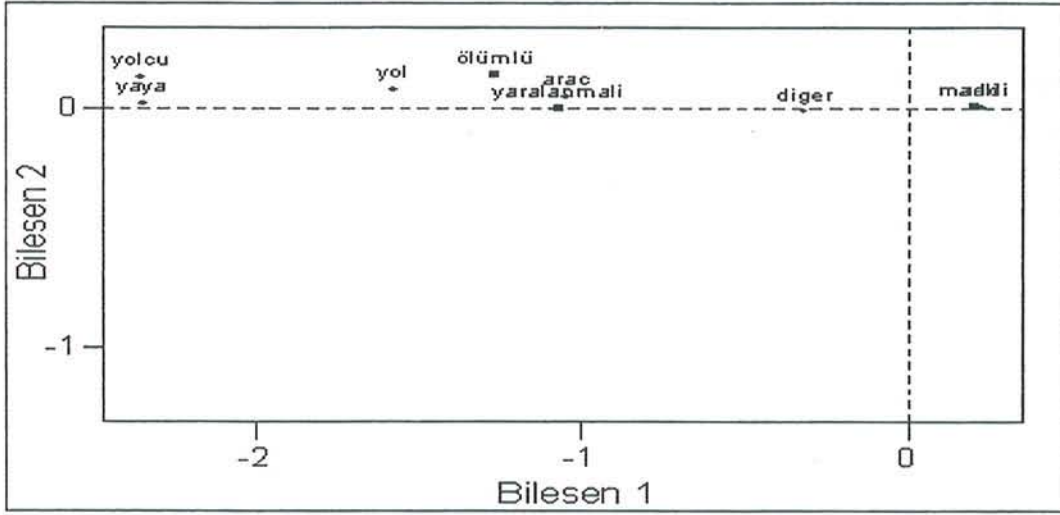
**Tablo 3.** Satır, sütun profilleri ve satır, sütun olasılıkları

	Satır Profilleri			Sütun Profilleri			Satır Olasılıkları
	Ölümlü Kaza	Yara. Kaza	M.Has. Kaza	Ölümlü Kaza	Yara. Kaza	M. Has. Kaza	
S. Asli	0,003	0,072	0,925	0,31	0,337	0,759	0,694
S.Diger	0,011	0,257	0,732	0,461	0,475	0,238	0,275
Yolcu	0,06	0,94	0	0,014	0,01	0	0,002
Yaya	0,051	0,949	0	0,181	0,152	0	0,024
Arac	0,03	0,501	0,469	0,014	0,011	0,002	0,003
Yol	0,042	0,679	0,279	0,02	0,015	0,001	0,003
Sütun Olasılıkları				0,007	0,148	0,845	

Tablo 3'de verilen satır ,sütun profilleri ile satır ve sütun olasılıklarından örnek olması amacı ile bir kaçının yorumunu yapalım. 0,003 satır profili bize sürücünün asli kusurlarından meydana gelen 1000 kazadan 3'ünün ölümlü kaza ile sonuçlandığını ifade eder. 0,337 sütun profili ise, meydana gelen 1000 yaralanmalı kazadan 337'sinin sürücünün asli kusurlarından kaynaklandığını gösterir. 0,694 satır olasılığı meydana gelen 1000 kazadan 694'ünün sürücünün asli kusurlarından kaynaklandığını gösterirken, 0,845 sütun olasılığı bize meydana gelen 1000 kazadan 845'inin maddi hasarlı kaza olduğunu göstermektedir. Uyum analizi sonrası elde edilen simetrik grafik şekil 1'de verilmiştir



**Şekil 1.** Tablo 1 için satır ve sütunların simetrik gösterimi



Şekil 1. Tablo 1 için satır ve sütunların simetrik gösterimi

Şekil 1 incelendiğinde, araçtan kaynaklanan kazaların yaralanmalı kazalarla birliktelik gösterdiği, maddi hasarlı kazaların sürücünün asli kusurları ile birliktelik gösterdiği söylenebilir. Kaza türünü en az etkileyen kusurlar ise, yolcu ve yaya kusurlarıdır.

Tablo 4. Yolun yüzeyine göre kazalar ve sonuçları

	Ölümlü Kaza	Yaralanmalı Kaza	Maddi Hasarlı Kaza	Toplam
Kuru	1855	43017	272194	317066
Islak	411	8978	73225	82614
Çamurlu	8	141	1251	1400
Karlı	27	520	4867	5414
Buzlu	29	664	3435	4128
Tozlu	29	376	508	913
Yağ ve Akaryakıt	154	4	51	99
Su Birikintili	442	5	104	333
Toplam	2368	53851	355912	412131

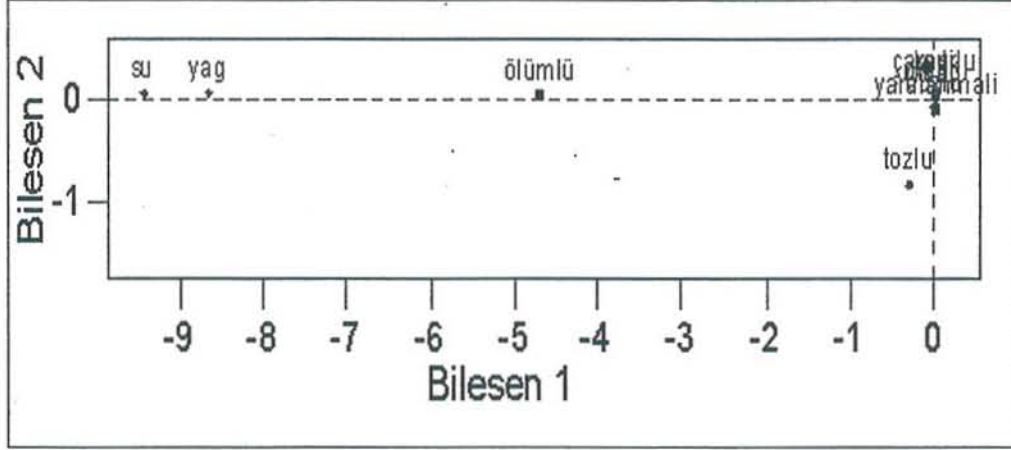
Tablo 4'de yolun yüzeyine göre kazalar ve sonuçları verilmiştir. Oluşturulan çarpımsal tablosuna uyum analizi uygulandıktan sonra elde edilen çarpımsal tablosu analizi Tablo 5'de gösterilmiştir.

Tablo 5. Tablo 4 için çarpımsal tablosu analizi

Bileşen	Varyans Ögesi	Oran	Birikimli
Bileşen 1	0,1572	0,982	0,982
Bileşen 2	0,0029	0,018	1
İz	0,16	$\chi^2$	65940,96

İz'in karekökünün ( $\sqrt{0,16}=0,4$ ) 0,2'den büyük olduğu görülmektedir. Buna göre kazanın olduğu yolun yüzeyi ile kaza türü arasındaki ilişki istatistiksel olarak anlamlıdır ve çarpımsal tablosuna uyum analizi uygulanması uygundur. Satır, sütun profilleri ile satır, sütun olasılıkları Tablo 6'da verilmiştir.

Şekil 2’de kazanın olduğu yolun yüzeyleri ile kaza sonuçlarının simetrik gösterimine yer verilmiştir. Şekil incelendiğinde, su birikintili , yağ ve akaryakıt birikintili yol yüzeylerinin kaza sonucunun türünü en az etkileyen düzeyler olduğu görülmektedir. Maddi hasarlı ve yaralanmalı kazalar aynı yol yüzeyi düzeylerinin etkisinde kaldığını ve yol yüzeylerinin ölümlü kazaların açıklanmasında fazla etkisi olmadığı söylenebilir.



Şekil 2. Tablo 4 için satır ve sütunların simetrik gösterimi

#### 4. SONUÇLAR

Trafik kazalarının uyum analizi ile incelendiği bu çalışma sonrasında, kaza sonucu ile kazaya sebep olan kusurlar ve kaza sonucu ile kazanın olduğu yol yüzeyi arasındaki ilişkilerin istatistiksel olarak anlamlı olduğu görülmüştür.

Uyum analizinde değişkenlerin düzeyleri arasındaki birliktelikler grafiksel olarak ortaya konmuştur. Maddi hasarlı ve yaralanmalı kazaların kuru, ıslak, çamurlu, karlı ve buzlu yol yüzeyleri ile birliktelik gösterdiği görülmüştür. Diğer çarpımsal tablosunun uyum analizi sonucunda, araçtan kaynaklanan kazaların yaralanmalı kazalarla, maddi hasarlı kazaların ise sürücünün asli kusurları ile birliktelik gösterdiği tespit edilmiştir.

#### KAYNAKÇA

- ANNA TORES & MICHAEL GREENACRE, *Measuring Asymmetries in Brand Associations Using Correspondence Analysis*, Journal of Economic Literature classification: C19, C88.
- BENDIXEN, M. T. (1995), *Compositional Perceptual Mapping Using Chi-squared Trees Analysis and Correspondence Analysis*, Journal of Marketing Management, Vol 11, No. 6, syf.571-581.
- BENZËCRI J. P. (1992), *Correspondence Analysis Handbook*, Marcel Dekker New York, Inc.
- Ç. A. SEYFULLAHOĞULLARI,(1996), *En Çok İşlem Hacmine Sahip On Şirketin Getiri ve Risklerine Göre Uygunluk Analizi İle Sınıflanması*,Ulusal Ekonometri ve İstatistik Sempozyumu Bildirileri, syf:511-21.

## KAYNAKÇA

- ANNA TORES & MICHAEL GREENACRE, *Measuring Asymmetries in Brand Associations Using Correspondence Analysis*, Journal of Economic Literature classification: C19, C88.
- BENDIXEN, M. T. (1995), *Compositional Perceptual Mapping Using Chi-squared Trees Analysis and Correspondence Analysis*, Journal of Marketing Management, Vol 11, No. 6, syf.571-581.
- BENZÈCRI J. P. (1992), *Correspondence Analysis Handbook*, Marcel Dekker New York, Inc.
- Ç. A. SEYFULLAHOĞULLARI,(1996), *En Çok İşlem Hacmine Sahip On Şirketin Getiri ve Risklerine Göre Uygunluk Analizi İle Sınıflanması*,Ulusal Ekonometri ve İstatistik Sempozyumu.Bildirileri, syf:511-21.
- GIFI A. (1990), *Nonlinear Multivariate Analysis*, New York, John Wiley&Sons.
- GREENACRE, M.J. (1984), *Theory and Applications of Correspondence Analysis*, Academic Press, London.
- GREENACRE M., BLASIUS J. (1994), *Correspondence Analysis in the Social Sciences*, Academic Press, London
- HOFFMAN. D. L. and FRANKE. G. R. (1986), *Correspondence Analysis:Graphical Representation of Categorical Data in Marketing Research*, Journal of Marketing Research, Vol. 13, syf:213-227.
- KURT FELLEBERG, NICOLE C. HAUSER, BENEDIKT BRORS, ALBERT NEUTZNER, JORG D. HOHEISEL and MARTIN VINGRON, (2001), *Correspondence analysis applied to microarray data*, PNAS Vol.98, No.19, syf:10781-10786.
- NURAN BAYRAM, MUSTAFA AYTAÇ, (2002), *Çoklu Karşılık Getirme Analizi İle Akademisyenlerin Akademik Niteliğinin Değerlendirilmesi*, İstatistik Araştırma Dergisi, Cilt 01 No 02, Özel Sayı, syf:39-54.
- ÖZDAMAR, K. (1999), *Paket Programlarla İstatistiksel Veri Analizi II*, 2. Baskı, Kaan Kitabevi, Eskişehir.
- PRESS W.H., TEUKOLSKY S. A., VETTERLING W. T., FLONNERY B.P., (1995) , *Numerical Recipes*, Cambridge University Press, USA
- STEN- ERIK CLAUSEN (1998), *Applied Correspondence Analysis*, Sage University Paper.
- Trafik İstatistik Yıllığı 2001, Emniyet Genel Müdürlüğü Trafik Hizmetleri Başkanlığı Yayını, Ankara



## THE USAGE OF CORRESPONDENCE ANALYSIS IN INVESTIGATING TRAFFIC ACCIDENTS

### ABSTRACT

*Correspondence analysis is a statistical technique based on geometric foundations that provides a graphical demonstration of multi-dimensional or two dimensional contingency tables in less dimensional plane handed by measuring correspondence between rows and columns. It is one of the most popular techniques that is frequently used in categorical data analysis for the aim of providing visual presentation between levels relating to variable groups. Specially, correspondence analysis that is frequently encountered in usage of sample surveys will be used in investigating traffic accidents in this study.*

**Key Words :** *Contingency Tables, Correspondence Analysis, SVD, Traffic Accidents.*