

KATEGORİK VERİLER İÇİN LOGARİTMİK DOĞRUSAL MODELLER VE GÖÇ İSTATİSTİKLERİ ÜZERİNE BİR UYGULAMA*

Sinan METE**
Aydın ÜNSAL***

ÖZET

İki yönlü olumsuzluk tablolarında istatistiki çıkarsamalar için Pearson'un ki-kare istatistiği yeterli olmakta ancak daha büyük boyutlu olumsuzluk tablolarında bu istatistik kullanılmamaktadır. Üç veya daha çok boyutlu olumsuzluk tablolarında ilişki yapılarının belirlenmesi için değişkenler arasında bağımlı – bağımsız ayrımı yapmayan “Logaritmik Doğrusal Modeller” kullanılabilir. Logaritmik Doğrusal Modeller yardımıyla daha çok değişken arasındaki etkileşimler sorgulanabilmektedir. Bu çalışmada Logaritmik Doğrusal Modeller incelenmiş ve göç istatistikleri üzerine bir uygulama yapılmıştır.

Anahtar Kelimeler: Kategorik Veri, Olumsuzluk Tabloları, Logaritmik Doğrusal Modeller

ABSTRACT

In two-ways contingency tables, Pearson's chi-square statistics is enough for statistical inferences but can't used in bigger contingency tables. For determining association structures in three-ways or bigger contingency tables, “Log-Linear Models” can be used, which doesn't discriminate variables as dependent or independent. Interactions between many variables can be interrogated with log-linear models. In this study, “Log-Linear Models” are investigated and an application on migration statistics had done.

Keywords: Categorical Data, Contingency Tables, Log-Linear Models

1. GİRİŞ

Sosyal bilimlerde yapılan çalışmalarda değişkenlerin niteliği, istatistiki çıkarsamaları ve analizleri çoğu zaman sınırlandırmaktadır. Nitel değişkenlerin kullanılmasının zorunlu olduğu bu dallarda veriler, parametrik teknik varsayımlarını çoğunlukla sağlamadığı için parametrik testlerle analiz edilemez.

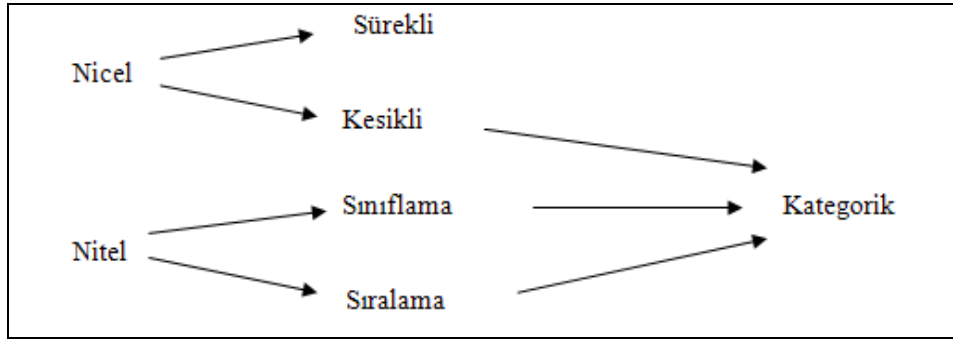
*Bu çalışmada, 2009 yılında Gazi Üniversitesi SBE Ekonometri ABD.'da Prof. Dr. Aydın ÜNSAL danışmanlığında hazırlanan ve kabul edilen “Kategorik Veri Analizi Yöntemleri ve Uygulamalar” başlıklı doktora tezinin ilgili bölümünden yararlanılmıştır.

**Arş.Gör.Dr., Aksaray Üniversitesi. İ.İ.B.F. İşletme Bölümü, Sayısal Yöntemler ABD, sinanmete@hotmail.com

***Prof. Dr., Gazi Üniversitesi İ.İ.B.F. Ekonometri Bölümü, aunsal@gazi.edu.tr

Değişkenlerin ölçümleri III. (eşit aralıklı) ve IV. (oranlı) ölçüm düzeylerinde yapıldığında analizlerde parametrik teknikler kullanılırken, düşük ölçüm düzeyleri olarak tanımlanan sınıflama (nominal) ve sıralama (ordinal) ölçüm düzeyleri söz konusu olduğunda parametrik olmayan tekniklerin kullanılması zorunlu olmaktadır.

Sınıflama veya sıralama ölçüm düzeyleriyle ölçülebilen veriler “Kategorik Veri” olarak adlandırılmaktadır, ancak küçük doğal sayı değerleri alan eşit aralıklı ölçme düzeyinde ölçülmüş veriler için de bazı durumlarda Kategorik Veri tanımı yapıp, parametrik olmayan teknikler kullanılır.



Şekil 1: Kategorik Veri Tanımı

Kaynak: Powers ve Xie, 2000;7.

Kategorik verilerin analizinde ilk aşama, veriler en yaygın ifade biçimi olan çapraz-sınıflandırılmış (kontenjans) tablo üzerinde toplanmalıdır. Kategorik değişken sayısının iki olduğu durumda, değişkenlerden biri satırda, diğeri ise sütunda olmak üzere iki-yönlü çapraz sınıflandırma tablosu oluşturulur.

İki değişkenden oluşan iki yönlü çapraz sınıflandırma tablosundan hareketle, iki değişken arasındaki ilişkiyi test etmek için, Pearson’un χ^2 istatistiği veya ilk defa Wilks tarafından hesaplanan G^2 olabilirlik oran istatistiği kullanılır.

Yalnızca iki değişken arasında ilişki olup olmadığını araştıran Pearson χ^2 test istatistiğinde n_{ij} gözle frekansları ile tahmin edilen f_{ij} beklenen frekansları arasındaki fark dikkate alınır ve beklenen değerler hesaplanırken değişkenlerin birbirinden bağımsız olduğu varsayılır.

$$\chi^2 = \sum \frac{(n_{ij} - f_{ij})^2}{f_{ij}}$$

Asimptotik olarak χ^2 dağılımı gösteren ve tam bölünebilme özelliğine sahip olduğu için de Pearson χ^2 test istatistiğine karşı üstünlük sağlayan G^2 olabirlik oran istatistiği de

$$G^2 = 2 \sum n_{ij} \ln \left(\frac{n_{ij}}{f_{ij}} \right)$$

olarak ifade edilebilir.

Sadece iki kategorik değişkenin bulunduğu iki-yönlü tabloların analizi Pearson χ^2 veya G^2 olabirlik oran istatistiği kullanılarak yapılırken, iki veya ikiden fazla değişkenden oluşan çok-yönlü tabloların analizi için logaritmik doğrusal modeller kullanılır.

2. LOGARİTMİK DOĞRUSAL MODELLER

İki kategorik değişkenden oluşan iki yönlü çapraz sınıflandırılmış tablolarda değişkenler arası ilişki Pearson χ^2 veya G^2 olabirlik oran istatistikleri yardımıyla test edilmekte iken, üç veya daha fazla değişkenden oluşan çok-boyutlu tabloların analizi için genellikle iki standart yol izlenmiştir (Bircan ve Yıldız, 1993; 24);

1. Yol: Çok boyutlu tablodan, mümkün olan iki-boyutlu marjinal toplamları elde ederek ayrı ayrı analiz etmek.

Bu metod, değişkenler arasında ilişki hakkında belirli bir doğrulukta bilgi sağlamakla birlikte, aşağıdaki mahsurları da taşımaktadır:

- Diğer değişkenler dikkate alındığında, ilişkili kategorik değişken çiftleri arasında marjinal ilişkiyi tam ayıramaz.
- Değişkenler arasındaki üç faktör ve daha yüksek dereceli interaksiyonların varlığını ihmal eder.
- Tüm ikili ilişkilerin aynı zamanda denemelerine müsaade etmez.

- **2. Yol:** Varyans analizi gibi, bilinen bazı parametrik metodları uygulamak.

Bu yolun rahatlıkla uygulanabilmesi varyans analizi varsayımlarından; x_{ij} gözlemlerinin varyansları homojen, gözlemlerin normal dağılışa uygun ve aynı zamanda toplanabilirlik varsayımlarının geçerli olması gerekmektedir.

Çapraz sınıflandırılmış tablolardaki değişkenlere ait veriler, multinomial, binom veya poisson gibi dağılımlara uygun olduğu için, varsayımların çoğunu yerine getiremezler. Bu varsayımları sağlayacak bazı transformasyonlar akla

gelmekle birlikte, yapılan bu transformasyonlar şartların tümünü sağlamayı başaramamaktadır.

Ki-Kare tekniği sadece iki yönlü çapraz tablolarda kullanılmaktadır. Ki-Kare tekniğinin üç veya daha çok yönlü tablolarda etkisiz kalmasından dolayı çok yönlü tabloların analizinde logaritmik doğrusal modeller kullanılmaktadır.

Nitel değişkenlerden oluşan çapraz tabloların çözümlenmesinde logaritmik doğrusal modellerin kullanımı üç amaca hizmet eder (Uygun, 1990;290);

- Değişkenlerin oluşturduğu bileşik dağılımı test etmek;
- Değişkenlerin birbirlerine bağımlı olup olmadığını test etmek,
- Değişkenler arasındaki ilişkiyi sebep-sonuç ilişkisine dayandırmaksızın test etmek.

Değişkenler arasındaki ilişki sebep-sonuç ilişkisine dayandırılarak test edilmek istenirse logaritmik doğrusal modellerin daha spesifik bir görünümü olan logit model ortaya çıkmaktadır (Acar, 2000; 784).

Ki-Kare tekniği ile değişkenler arasında ilişki test edilirken parametre tahmini yapılmaz, logaritmik doğrusal modellerde ise bir model oluşturulur ve modeldeki parametrelerin tahmini yapılır. Gözlemlerle ilgili kuramsal çerçeveyi tanımlayan “model”den elde edilen parametreler, değişkenlerin bileşimini ya da etkileşimini tanımlamaktadır (Mete, 2009; 3).

Log-linear modeller kategorik değişkenlerin analizinde sıklıkla kullanılan bir yöntemdir. Daha çok iki veya daha fazla kategorili ikiden fazla değişkenin (çok yönlü kontenjans tablolarının) analizinde kullanıldığından “Çok-Yönlü Frekans Analizi” olarak da adlandırılmaktadır. “Log-Linear Modeller”, kontenjans tablolarının hücrelerinde yer alan frekansların, değişkenlerin kategorilerine ne kadar bağlı olduğunu ortaya çıkaran, kategorik değişkenler arasındaki ilişkileri ve etkileşimleri inceleyen modellerdir (Çılan, 2009;153).

Log-linear modellerde parametrelerin tahmin edicilerini bulabilmek için odds oranlarından yararlanır. Odds, herhangi bir değişkenin bir seçeneğinde bulunan frekansın aynı değişkenin diğer seçeneğindeki frekansına oranı şeklinde tanımlandığında, bitişik satırlardan (sıra i ve sıra $i+1$) ve bitişik sütunlardan (sütun j ve $j+1$) düzenlenen ve beklenen frekansları esas alan 2×2 alt tablolar için uygun odds oranı ($i=1,2,\dots,I-1$ ve $j=1,2,\dots,J-1$) θ_{ij} ile gösterildiğinde

$$\theta_{ij} = \frac{(F_{ij} F_{i+1,j+1})}{(F_{i,j+1} F_{i+1,j})}$$

olarak hesaplanır. Denklemden tanımlanan θ_{ij} $(I-1)(J-1)$ odds-oranları ile, $I \times J$ tablosundan düzenlenen herhangi diğer 2×2 alt tablolara ait olan odds-oranları belirlenebilir (Bayram, 2000; 263).

Logaritmik doğrusal model bağımlı ve bağımsız değişken ayrımı yapmaz. İncelenecek kriter beklenen hücre frekanslarıdır (Green vd., 1977; 54).

Logaritmik doğrusal modellerde parametre sayısı, değişken sayısı (d)'ye bağlı olarak 2^d 'den hesaplanır. İki boyutlu modeller için (2^2) dört adet parametre bulunurken, üç boyutlu modelde parametre sayısı (2^3) sekiz adet olur.

2.1. İki Boyutlu Tabloların Analizi

Her biri iki düzeye sahip iki kategorik değişken değerlerinin biri satırda, diğeri sütunda olmak üzere 2×2 boyutlu bir çapraz sınıflandırılmış tabloda gösterildiğini düşünelim. Bu tablo için logaritmik doğrusal model

$$\ln(F_{ij}) = \mu + \lambda_i^A + \lambda_j^B + \lambda_{ij}^{AB} \quad (1)$$

biçimindedir. Doymuş model olarak adlandırılan (1) numaralı model hem tekli hem de ikili etkileşimleri içermektedir;

$\ln(F_{ij})$: Tablodaki ij gözesi için beklenen frekansların doğal logaritması,

μ : Beklenen frekansların doğal logaritma değerlerinin ortalaması,

λ_i^A : A değişkeninin ana etkisi,

λ_j^B : B değişkeninin ana etkisi,

λ_{ij}^{AB} : A ve B değişkenlerinin etkileşim etkisi.

λ_{ij}^{AB} parametresinin 0 değerini alması, A değişkeni ile B değişkeninin bağımsız olmasını ifade eder ve (1) numaralı model,

$$\ln(F_{ij}) = \mu + \lambda_i^A + \lambda_j^B \quad (2)$$

biçiminde, doymamış bir modele dönüşür.

Logaritmik doğrusal modellerde yüksek dereceli bir terimin modelde bulunduğu zaman, daha düşük dereceli tüm terimlerin de modelde yer aldığı türde ilişkiler "hiyerarşik" olarak nitelendirilir. Hiyerarşik bir logaritmik doğrusal modelde, bir λ parametresi sıfıra eşit ise, daha yüksek dereceli tüm λ

parametreleri de sifira eşit olur. Eğer, λ parametresi sıfırdan farklı ise, daha düşük dereceli tüm parametreler de sıfırdan farklıdır. Örneğin;

$$\text{Ln}(F_{ij}) = \mu + \lambda_i^A + \lambda_{ij}^{AB} \quad (3)$$

modeli hiyerarşik olmayan bir modeldir, çünkü λ_{ij}^{AB} parametresi modelde bulunmasına rağmen λ_j^B ana etkileşim terimi modelde yer almamaktadır.

2.2. Üç Boyutlu Tabloların Analizi

Üç değişkenin yer aldığı bir çapraz sınıflandırma tablosu için doymuş logaritmik doğrusal model;

$$\text{Ln}F_{ijk} = \lambda + \lambda_i^X + \lambda_j^Y + \lambda_k^Z + \lambda_{ij}^{XY} + \lambda_{ik}^{XZ} + \lambda_{jk}^{YZ} + \lambda_{ijk}^{XYZ} \quad (4)$$

dir.

X, Y, Z gibi üç değişken içeren çapraz sınıflandırma tablosundan hiyerarşik olarak elde edilebilecek logaritmik doğrusal modeller aşağıdaki gibi sıralanabilir;

Tam bağımsızlık modeli;

$$\text{Ln}F_{ijk} = \lambda + \lambda_i^X + \lambda_j^Y + \lambda_k^Z \quad (5)$$

kısmi bağımsızlık içeren modeller;

$$\text{Ln}F_{ijk} = \lambda + \lambda_i^X + \lambda_j^Y + \lambda_k^Z + \lambda_{ij}^{XY}$$

$$\text{Ln}F_{ijk} = \lambda + \lambda_i^X + \lambda_j^Y + \lambda_k^Z + \lambda_{jk}^{YZ}$$

$$\text{Ln}F_{ijk} = \lambda + \lambda_i^X + \lambda_j^Y + \lambda_k^Z + \lambda_{ik}^{XZ}$$

koşullu bağımsızlık içeren modeller;

$$\text{Ln}F_{ijk} = \lambda + \lambda_i^X + \lambda_j^Y + \lambda_k^Z + \lambda_{ij}^{XY} + \lambda_{ik}^{XZ}$$

$$\text{Ln}F_{ijk} = \lambda + \lambda_i^X + \lambda_j^Y + \lambda_k^Z + \lambda_{ij}^{XY} + \lambda_{jk}^{YZ}$$

$$\text{Ln}F_{ijk} = \lambda + \lambda_i^X + \lambda_j^Y + \lambda_k^Z + \lambda_{ik}^{XZ} + \lambda_{jk}^{YZ}$$

tüm ikili etkileşimleri içeren model ise;

$$\text{Ln}F_{ijk} = \lambda + \lambda_i^X + \lambda_j^Y + \lambda_k^Z + \lambda_{ij}^{XY} + \lambda_{ik}^{XZ} + \lambda_{jk}^{YZ}$$

dir.

Bu hiyerarşik modellerde, λ genel ortalamayı, λ_i^X , λ_j^Y , λ_k^Z ana etkileri, λ_{ij}^{XY} , λ_{ik}^{XZ} , λ_{jk}^{YZ} ikili etkileşimleri ve λ_{ijk}^{XYZ} de üçlü etkileşimi gösterir.

İstatistiksel paket programları hiyerarşik olmayan modelleri test edememektedir. Çünkü hiyerarşik olmayan modeller arasından seçim için bir istatistiksel süreç sağlanamamaktadır. Bunun sebebi ise bir etkileşim teriminin anlamlı olmadığı durumlarda, daha yüksek dereceli etkileşimlerin anlamlı olup olmadığının yorumunun anlam taşımamasındandır. (Oğuzlar, 2004;237)

Üç yönlü hiyerarşik bir logaritmik doğrusal model için test edilmesi gereken başlıca hipotezler şunlardır;

$$H_{XYZ} : \lambda_{ijk}^{XYZ} = 0 \quad (6)$$

$$H_{XY} : \lambda_{ijk}^{XYZ} = 0 \quad \lambda_{ij}^{XY} = 0 \quad (7)$$

$$H_{XZ} : \lambda_{ijk}^{XYZ} = 0, \quad \lambda_{ik}^{XZ} = 0 \quad (8)$$

$$H_{YZ} : \lambda_{ijk}^{XYZ} = 0, \quad \lambda_{jk}^{YZ} = 0 \quad (9)$$

$$H_X : \lambda_{ijk}^{XYZ} = 0, \quad \lambda_{ij}^{XY} = \lambda_{ik}^{XZ} = \lambda_{jk}^{YZ} = 0, \quad \lambda_i^X = 0 \quad (10)$$

$$H_Y : \lambda_{ijk}^{XYZ} = 0, \quad \lambda_{ij}^{XY} = \lambda_{ik}^{XZ} = \lambda_{jk}^{YZ} = 0, \quad \lambda_j^Y = 0 \quad (11)$$

$$H_Z : \lambda_{ijk}^{XYZ} = 0, \quad \lambda_{ij}^{XY} = \lambda_{ik}^{XZ} = \lambda_{jk}^{YZ} = 0, \quad \lambda_k^Z = 0 \quad (12)$$

$$H_0 : \lambda_{ijk}^{XYZ} = 0, \quad \lambda_{ij}^{XY} = \lambda_{ik}^{XZ} = \lambda_{jk}^{YZ} = 0, \quad \lambda_i^X = \lambda_j^Y = \lambda_k^Z = 0 \quad (13)$$

Logaritmik Doğrusal modellerde test edilen değişken ve oluşturulan hipotez sayısı çok fazladır. Bu nedenle örneklem hacminin küçük seçilmesi sonuç elde etmeyi güçleştireceğinden büyük hacimli örneklemle çalışılması gerekmektedir. Bu doğrultuda örneklem hacminin tablodaki hücre sayısının en az 5 katı olması önerilmektedir (Akgül, 1997; 408).

2.3. En İyi Modelin Seçimi

Etkileşimlerin test edilmesinde iki farklı yaklaşım söz konusudur. Örneğin üç değişken içeren üç yönlü çapraz sınıflandırılmış tablolar için (4) numaralı doymuş modelden başlayarak, kabul edilebilir düzeye gelinceye kadar yüksek dereceden etkileşim terimlerinin modelden atılması ile en uygun modele ulaşılabilir. Diğer yaklaşımda ise, başlangıç modeli (5) numaralı tam bağımsızlık modelidir. Bu modele kabul edilebilir düzeye ulaşılan kadar etkileşim terimleri eklenerek en uygun model elde edilebilir.

Birden fazla modelin anlamlı çıkması durumunda, hangi logaritmik doğrusal modelin seçilen değişkenler arasındaki ilişkiyi doğru ifade ettiğini belirlemede uyum iyiliği testleri kullanılır. χ^2 ve G^2 olabirlik oran test

istatistikleri en çok kullanılan uyum iyiliği testleridir. G^2 olabilirlik oran test istatistiği,

$$G^2 = 2 \sum_{i=1}^I \sum_{j=1}^J \sum_{k=1}^K f_{ijk} \ln \left(f_{ijk} / \hat{F}_{ijk} \right)$$

gibi hesaplanabilir.

M_1 ve M_2 iki model olsun. M_1 , M_2 'ye göre daha basit bir model olmak şartıyla iki modelden hangisinin uygun log-lineer model olduğuna G^2 olabilirlik oran test istatistikleri arasındaki fark alınarak karar verilir. Koşullu olabilirlik oran test istatistiği adı verilen bu test istatistiği,

$$G^2(M_1 / M_2) = G^2(M_1) - G^2(M_2)$$

düzenlenen sıfır hipotezi altında $v_1 - v_2$ serbestlik derecesi ile asimptotik olarak χ^2 dağılır.

Üç yönlü tablolar için yorumlanan logaritmik doğrusal modeller daha fazla değişken içeren tablolar için de genişletilebilir. Ancak değişken sayısının artması, ilişki ve etkileşim terimlerinin yorumlanmasını, test edilmesi gereken hipotezlerin oluşturulmasını ve en iyi model seçim sürecini daha karmaşık ve zor yorumlanabilir bir hale sokmaktadır.

3. UYGULAMA

Çalışmada, 22 Ekim 2000 tarihinde Türkiye’de yapılan ondördüncü genel nüfus sayımının göç ile ilgili bilgileri kullanılmıştır. Türkiye İstatistik Kurumu’nun (TÜİK) yaptığı nüfus sayımının göç ile ilgili detayları 2005 yılında, ” Genel Nüfus Sayımı 2000 Göç İstatistikleri” adlı bir kaynaktan toplanarak basılmıştır.

Sayım günündeki daimi ikametgah yeri ile sayımdan beş yıl önceki daimi ikametgah yeri farklı olan kişiler, göç eden nüfus olarak tanımlanmıştır. Bu tanım gereği, göç eden nüfus, 6 ve daha yukarı yaştaki nüfusu kapsamaktadır. (TÜİK, 2005;27)

Çalışmadaki veriler, 1985-2000 yılları arasındaki göç istatistiklerinden alınmıştır. Uygulamadaki değişkenler (göç türü, cinsiyet, okuma yazma) ve bu değişkenlerin kategorileri aşağıdaki gibidir.

<i>Göç Türü</i>	<i>Cinsiyet</i>	<i>Okuma Yazma</i>
Şehirden Şehire	Erkek	Okuma Yazma Bilen
Köyden Şehire	Bayan	Okuma Yazma Bilmeyen
Şehirden Köye		
Köyden Köye		

1985-2000 yıllarında toplam 6.579.460 kişi (6 yaş ve yukarı) göç etmiştir. Göç edenlerin; Göç Türü, Cinsiyet ve Okuma Yazma değişkenlerine göre yerleştirildiği üç-yönlü çapraz tablo 0.05 yanılma düzeyinde, SPSS for Windows 15.0 paket programı yardımıyla Logaritmik Doğrusal modellerle analiz edilmiştir.

Öncelikle üç yönlü tabloda ana etki, ikinci dereceden etkileşim ve üçüncü dereceden etkileşim terimlerinin anlamlılığı incelenmiştir.

Tablo 1: Göç Türü, Cinsiyet ve Okuma Yazma Değişkenleri İçin K – Yönlü Etkileşim Özet Tablosu

	K	s.d.	Olabilirlik Oran		Pearson		İterasyon Sayısı
			Ki-Kare	p değeri	Ki-Kare	p değeri	
K-Yönlü ve Daha Yüksek Etkileşimler	1	15	9963726,0	,000	12814477	,000	0
	2	10	228941,782	,000	248831,30	,000	2
	3	3	271,708	,000	270,857	,000	3
K-Yönlü Etkileşimler	1	5	9734784,2	,000	12565646	,000	0
	2	7	228670,074	,000	248560,44	,000	0
	3	3	271,708	,000	270,857	,000	0

s.d.= Serbestlik Derecesi

Tablo 1’deki ilk kısmın birinci satırı, “bütün birinci ve daha üst dereceli etkilerin 0’a eşit olduğu”, yani ana etkiler ve daha büyük etkilerin (ikili ve üçlü), sıfır olduğu hipotezini,

$$H_0 : \lambda_{ijk}^{XYZ} = 0, \quad \lambda_{ij}^{XY} = \lambda_{ik}^{XZ} = \lambda_{jk}^{YZ} = 0, \quad \lambda_i^X = \lambda_j^Y = \lambda_k^Z = 0$$

test eder. Gerek olabilirlik oran, gerekse Pearson Ki-Kare test istatistiklerinin yüksek, ve p değerinin 0.000 olması sonucunda ($p = 0.000 < 0.05$), sıfır hipotezi reddedilir.

İkinci satırda ise, “bütün ikinci dereceden ve üçüncü dereceden etkilerin sıfıra eşit olduğu” hipotezi,

$$H_{XY} : \lambda_{ijk}^{XYZ} = 0, \quad \lambda_{ij}^{XY} = \lambda_{ik}^{XZ} = \lambda_{jk}^{YZ} = 0$$

yüksek olabilirlik oran ve Pearson Ki-Kare test istatistikleri ve 0.000 olan p değeri ($p = 0.000 < 0.05$) ile reddedilir.

Tablo 1.’in ilk kısmının son satırında ise,

$$H_{XYZ} : \lambda_{ijk}^{XYZ} = 0$$

hipotezi, yani “üçlü etkileşimin sıfıra eşit olduğu” hipotezi, 0.000 olan p değeri ($p = 0.000 < 0.05$) ile reddedilir.

Tablo 1’in ikinci kısmı ise, sırasıyla; “birinci, ikinci ve üçüncü dereceden etkilerin sıfıra eşit olduğu” hipotezlerini test eder. Bu hipotezler, test istatistik sonuçlarının yüksek olması ve 0.000 olan p değeri ($p = 0.000 < 0.05$) ile reddedilir.

Modelde yer alacak ikinci dereceden etkileşim parametrelerini tespit etmek için aşağıdaki Tablo 2’den kısmi ilişki testleri incelenebilir.

Tablo 2’deki ikinci dereceden etkileşim parametreleri incelendiğinde, cinsiyet-egitim, cinsiyet-göç türü ve eğitim-göç türü değişkenleri için kısmi ilişki parametreleri istatistiksel olarak anlamlı bulunmuştur.

Tablo 2: Göç Türü, Cinsiyet ve Okuma Yazma Değişkenlerinin Etkileşim Düzeylerine Ait Ki-Kare ve Olasılık Değerleri

Etki	Serbestlik Derecesi	Kısmi Ki-Kare	p değeri	İterasyon Sayısı
cins*egitim	1	166993,536	,000	2
cins*göçtür	3	1317,172	,000	2
egitim*göçtür	3	61458,872	,000	2
cins	1	52976,615	,000	2
egitim	1	5737872,6	,000	2
göçtür	3	3943933,4	,000	2

Tablo 2’deki ana etkileri gösteren cinsiyet, eğitim ve göç türü değişkenlerine ait parametre tahminleri 0.05 yanılma düzeyinde istatistiksel olarak anlamlı bulunmuştur.

Tablo 1 ve Tablo 2’den elde edilen sonuçlar, aşağıdaki model parametre tahminlerinin yer aldığı Tablo 3’ten de görülebilir.

Tablo 3: Göç Türü, Cinsiyet ve Okuma Yazma Değişkenlerinin Parametre Tahminleri

Etki	Parametre	Katsayı	Std. Hata	Z	p değeri	95% Güven Aralığı	
						Alt Sınır	Üst Sınır
cins*egitim*göçtür	1	,008	,001	6,139	,000	,006	,011
	2	-,023	,002	-14,809	,000	-,027	-,020
	3	-,004	,002	-2,497	,013	-,007	-,001
cins*egitim	1	-,326	,001	-322,627	,000	-,328	-,324
cins*göçtür	1	-,006	,001	-4,603	,000	-,009	-,004
	2	,004	,002	2,482	,013	,001	,007
	3	,001	,002	,575	,565	-,002	,004
egitim*göçtür	1	-,296	,001	-216,133	,000	-,299	-,294
	2	,087	,002	54,557	,000	,083	,090
	3	-,026	,002	-16,488	,000	-,030	-,023
cins	1	-,188	,001	-186,347	,000	-,190	-,187
egitim	1	-1,221	,001	-1206,735	,000	-1,223	-1,219
göçtür	1	,946	,001	689,622	,000	,943	,948
	2	,066	,002	41,755	,000	,063	,069
	3	,117	,002	73,223	,000	,114	,121

Tablo 3'te p değerlerinin tamamına yakını sıfır çıkmış ve hemen hemen değişkenlerin her düzeyinde etkileşimlerin sıfırdan farklı olduğu hipotezi reddedilmiştir.

4. SONUÇ

Kategorik verilerle çalışılan alanlarda, veriler parametrik test varsayımlarını geçemediği için parametrik analizler yapılamamaktadır. Kategorik iki değişken arasındaki ilişki olup olmadığı ki-kare analizi ile test edilebilirken, ikiden fazla değişken barındıran çok-yönlü tabloların analizinde ki-kare analizi kullanılamamaktadır.

Kategorik verilerden oluşan çok-yönlü tabloların analizi için logaritmik doğrusal modeller, bağımlı-bağımsız ayrımı yapmaksızın, değişkenler arasında ilişki olup olmadığını test edebilmektedir. Logaritmik doğrusal modellerin spesifik bir türü olan Logit modeller değişkenler arasında bağımlı-bağımsız ayrımı yaparak, ilişkileri test etmektedir.

Çalışmada, iki-yönlü ve üç-yönlü çapraz tabloların logaritmik doğrusal modellerle analiz edilmesi teorik olarak açıklanmış ve 1985-2000 yılları arasındaki göç istatistiklerinden derlenmiş bir üç-yönlü tabloya uygulanmıştır.

Uygulamada ele aldığımız üç-yönlü tablonun, logaritmik doğrusal modellerle analiz edilmesi sonucunda gerek üçüncü dereceden etkileşim ve gerekse ikinci dereceden etkileşimlerin sıfırdan farklı olduğu hipotezleri reddedilmiş ve doymuş bir model elde edilmiştir.

5. KAYNAKLAR

- ACAR, Fatma. (1999). *Çapraz Tabloların Çözümlemesinde Logaritmik Doğrusal Modellerin Kullanımı*, IV. Ulusal Ekonometri ve İstatistik Sempozyumu Bildirileri, 14-16 Mayıs, Antalya, (s:783-799).
- AKGÜL, Aziz. (1997). *Tıbbi Araştırmalarda İstatistiksel Analiz Teknikleri - SPSS Uygulamaları*. Yükseköğretim Kurulu Matbaası, Ankara.
- BAYRAM, Nuran. (2000). Kategorik Verilerin Log-Linear Modellerle Analizi. *Uludağ Üniversitesi İ.İ.B.F. Dergisi*, 19, 1-2, (s:261-267).
- BİRCAN, Hüdaverdi ve YILDIZ, Necati. (1993). Üç-Boyutlu Çapraz-Sınıflandırılmış Kategorik Verilere Log-Linear Modellerin Uydurulması, *Atatürk Ü. Ziraat Fakültesi Dergisi* 24 (2), (s:21-35).
- ÇİLAN, Çiğdem Arıçgil. (2009). *Sosyal Bilimlerde Kategorik Verilerle İlişki Analizi-Kontenjans Tabloları Analizi*, Pegem Akademi, Ankara.
- GREEN, P.E. vd. (1977). On the Analysis of Qualitative Data in Marketing Research, *Journal of Marketing Research*, Vol: XIV, (s: 52-59).
- METE, Sinan. (2009). *Kategorik Veri Analizi Yöntemleri ve Uygulamalar*, Gazi Üniversitesi Sosyal Bilimler Enstitüsü, Yayınlanmamış Doktora Tezi, Ankara.
- OĞUZLAR, Ayşe. (2004). Hiyerarşik Logaritmik Doğrusal Modeller Arasından En Uygun Modelin Seçimi, *Öneri*, C:6, S:21, (s:235-245).
- POWERS, Daniel A. ve XIE, Yu. (2000). *Statistical Methods For Categorical Data Analysis*, Academic Press. U.S.A.
- TÜİK. (2005). Yayın No:2976, *Genel Nüfus Sayımı 2000 Göç İstatistikleri*, TÜİK Matbaası, Aralık.
- UYGUN, Hamza. (1990). Çapraz Tabloların Çözümlemesi ve Log-Linear Modeller, *Hacettepe Üniversitesi İktisadi İdari Bilimler Fakültesi Dergisi*, Cilt:8, Sayı:1, (s:299-308).