Öneri.C.5.S.20.
Haziran 2003.ss.213-223.

LİSANSÜSTÜ EĞİTİME GİRİŞ SINAVI (LES) SONUÇLARININ ÜÇ YÖNLÜ ÇAPRAZ SINIFLANDIRMA TABLOSU İLE İNCELENMESİ

Dilek ALTAŞ[1], Esen ZEREN YILDIRIM[2]

[1]Marmara Üniversitesi, İ.İ.B.F., Ekonometri Bölümü, Yardımcı Doçent Dr.
[2]Marmara Üniversitesi, İ.İ.B.F., Ekonometri Bölümü, Araştırma Görevlisi

EXAMINING THE RESULTS OF LES WITH THREE-WAY CROSS CLASSIFIED TABLE

***Abstract:** While the examining of the relations between the categorical variables, the observation values are collected on a cross classified table. In the situation that there is two categorical variables, then a two dimensional table is prepared as if one variable is in the row, and the other is in the column. With respect to this table, in order to examine the relations between the variables, usually applied method is chi-square analysis.*

When the variables is more than two, the analysis of variables' cross classified tables as one within the other can not be done with the chi-square analysis. In this case, for the examining of the relation structures between the variables, log-linear analysis can be made use of. Log-linear analysis is a method that can analyze with the help of the models, both IX2 or 2XJ dimensional tables to which chi- square can be applied, and three or more dimensional tables to which chi-square is insufficient.

In this study, the results of the Entrance Examination to the Postgraduate Education (LES) carried out by The Student Selection and Placement Center (OSYM) are examined with the log-linear analysis by the classification on the three dimensional crosswise table. In the scope of the analysis, the factors, that is examined with their relation structures, are the variables of education level (bachelor's degree, master, doctor's degree), LES score (45 point and more or less 45 point) and the type of point (qualitative, quantitative and equal weight). At the end of the analysis, the relation structure between the in question variables is revealed and the findings are interpreted.

Keywords: Chi-Square Independence Test, Log-Linear Models, Odds Ratio, LES Scores.

LİSANSÜSTÜ EĞİTİME GİRİŞ SINAVI(LES) SONUÇLARININ ÜÇ YÖNLÜ ÇAPRAZ SINIFLANDIRMA TABLOSU İLE İNCELENMESİ

***Özet:** Kategorik değişkenler arasındaki ilişkiler incelenirken, gözlem değerleri bir çapraz sınıflandırma tablosu üzerinde toplanır. Kategorik değişken sayısının iki olması durumunda, değişkenlerden biri satırda, diğeri sütunda olacak şekilde iki yönlü bir çapraz tablo oluşturulur. Bu tablodan hareketle, değişkenler arasındaki ilişkileri incelemek amacıyla sıklıkla başvurulan yöntem, ki-kare analizidir.*

Ki-kare analizi ile ikiden çok sayıda değişkenin iç içe çapraz tablolarının analizi yapılamamaktadır. Bu durumda, değişkenler arasındaki ilişki yapısını incelemede loglineer analizden yararlanılabilir. Loglineer analiz, hem ki-karenin uygulanabildiği IX2 ya da 2XJ boyutlu tabloları, hem de ki-karenin yetersiz kaldığı üç ve daha büyük boyutlu tabloları modeller aracılığıyla analiz eden bir yöntemdir.

Bu çalışmada, Öğrenci Seçme ve Yerleştirme Merkezi (ÖSYM) tarafından uygulanan Lisansüstü Eğitime Giriş Sınavı (LES) sonuçları, üç yönlü bir çapraz tablo üzerinde sınıflandırılarak loglineer analizle incelenmiştir. Analiz kapsamında aralarındaki ilişki yapısı incelenen faktörler, öğrenim durumu (lisans, yüksek lisans, doktora mezunu olma), LES puanı (45 yada daha fazla, 45'den az) ve puan türü (sayısal, sözel, eşit ağırlıklı) değişkenleridir. Analiz sonucunda, söz konusu değişkenler arasındaki ilişki yapısı ortaya çıkarılmış, bulgular yorumlanmıştır.

Anahtar Kelimeler: Ki-Kare Bağımsızlık Testi, Loglineer Modeller, Odds Oranı, LES Sonuçları.

I. GİRİŞ

Öğrenci Seçme ve Yerleştirme Merkezi (ÖSYM) tarafından uygulanan Lisansüstü Eğitime Giriş Sınavı (LES), yüksek lisans ve doktora eğitimi almak isteyenler ile araştırma görevliliğine atanacaklar için önkoşul olarak getirilen merkezi bir sınavdır. Sınavda adaylara yöneltilen sorular, sayısal ve sözel olmak üzere iki bölümden oluşmakta, sınav sonuçları sayısal, sözel ve eşit ağırlıklı olmak üzere üç puan türünde ilan edilmektedir.

Lisansüstü eğitim görmek isteyen öğrenciler, başta sosyal bilimler, fen bilimleri ve sağlık bilimleri olmak üzere üniversitelerin çeşitli enstitülerine başvuruda bulunmaktadır. 2547 sayılı kanunun 35. maddesi uyarınca öğrencilerin, başvurdukları programın puan türünde 45 standart puandan az olmamak koşuluyla, ilgili üniversite yönetimi tarafından belirlenecek LES puanına sahip olmaları gerekmektedir. Öğrencilerin lisansüstü programlara yerleştirilmeleri, LES puanı, mezuniyet notu ve mülakat notunun ağırlıklı ortalaması dikkate alınarak yapılmaktadır. Bu puanların giriş notunun hesaplanmasındaki ağırlıkları, üniversite yönetimince belirlenmektedir.

Bu çalışmada LES sonuçları, üç yönlü bir çapraz tablo üzerinde sınıflandırılarak kategorik veri analizine imkan sağlayan loglineer analizle incelenmiştir. Analiz kapsamında aralarındaki ilişki yapısı incelenen faktörler, öğrenim durumu (lisans, yüksek lisans, doktora mezunu), LES puanı (45 ve üstü, 45'in altı) ve puan türü (sayısal, sözel, eşit ağırlıklı)dür. Elde edilen loglineer model ve parametre tahminlerine ilişkin bulgular, uygulama bölümünde sunulmuştur.

II. KATEGORİK VERİLERDE İLİŞKİ ANALİZİ

Kategorik değişkenler arasındaki ilişkileri incelemenin ilk aşaması, gözlem değerlerini çapraz sınıflandırma (kontenjans) tablosu üzerinde toplamaktır. Kategorik değişken sayısının iki olması durumunda, değişkenlerden biri satırda, diğeri sütunda olacak şekilde iki yönlü çapraz sınıflandırma tablosu oluşturulur. Değişken sayısının ikiden fazla olması durumunda ise, tablo genişletilerek üç yönlü, dört yönlü tablolar elde edilebilir.

Odds ve Odds Oranları

X ve Y değişkenleri, iki kategoriye sahip dikotom (binary) değişkenler olsun. Yaygın olarak dikotom değişkenlerin alabileceği değerler "başarılı" ve "başarısız" olarak sınıflandırılır. i'inci satırda Y için başarı olasılığı π_i, başarısızlık olasılığı (1-π_i) olsun. Bu ($\pi_i, 1-\pi_i$) koşullu olasılıkları kullanılarak, iki satırda başarı olasılıklarının farkı test edilebilir.

p_1 ve p_2 örneklemden bulunan başarı olasılıkları, n_1 ve n_2 örneklem büyüklükleri olmak üzere, $\pi_1 - \pi_2$ için güven aralığı aşağıdaki şekilde hesaplanabilmektedir.

$$(p_1-p_2) \pm z_{\alpha/2}\hat{\sigma}(p_1 - p_2),$$

Burada,

$$\hat{\sigma}(p_1 - p_2) = \sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}}$$

Başarı olasılıkları test edilirken, oranların 0'a ya da 1'e yakın olması durumunda oranların farkı yoluyla yapılan hipotez testi, yanıltıcı sonuçlar verebilir. Örneğin, 0,010 ve 0,001 arasındaki fark, 0,410 ve 0,401 arasındaki farka eşittir. Bu oranların iki ayrı ilacın kullanımı sonucunda yan etkiye maruz kalan hastaların oranları olduğu düşünülürse, 0,010 değeri 0,001'in 10 katıdır, buna karşın 0,410 ve 0,401 arasındaki fark çok küçük kalmaktadır. Bu sebeple, oranların farkı ($\pi_1 - \pi_2$) yerine oranların oranı $\left(\frac{\pi_1}{\pi_2}\right)$ daha iyi bir gösterge olacaktır. Bu değere *relatif risk* adı verilir [1].

Relatif riskten hareketle, çapraz sınıflandırma tablolarının analizinde sıkça kullanılan *odds* değeri tanımlanabilir. Odds değeri, bir sonucun gerçekleşme olasılığının gerçekleşmeme olasılığına oranıdır [1].

$$\text{odds} = \frac{\pi}{1-\pi}$$

Odds değeri negatif olmayan değerler alır. Odds=1 ise, başarı ve başarısızlık olasılıkları aynı demektir. Odds>1 ise, başarının gerçekleşme şansı başarısızlığa göre daha yüksektir. Odds<1 ise, tersi geçerlidir.

Yalnızca bir değişken için hesaplanan odds değerlerini, iki değişken için hesaplanan odds değerlerinden ayırmak amacıyla, marjinal odds ve koşullu odds ayrımı yapılır. Odds değeri, değişkenlerin her biri için satır ya da sütun toplamları dikkate alınarak bulunmuşsa *marjinal odds*, değişkenlerden birinin belirli bir kategorisinde diğer değişken için bulunmuşsa *koşullu odds* adını alır [2].

2x2'lik çapraz sınıflandırma tablolarında ilişki analizinde kullanılan ve bir çok model için de parametre kabul edilen bir diğer ölçü, *odds oranı*dır. Odds oranı, iki koşullu oddsun oranı biçiminde aşağıdaki gibi hesaplanır [3].

$$\theta = \frac{odds_1}{odds_2} = \frac{\pi_1/(1-\pi_1)}{\pi_2/(1-\pi_2)}$$

θ odds oranı bileşik olasılıklar cinsinden aşağıdaki gibi de yazılabilir. Elde edilen oran, yaygın olarak *çapraz çarpım oranı* adıyla bilinmektedir [3].

$$\theta = \frac{\pi_1/(1-\pi_1)}{\pi_2/(1-\pi_2)} = \frac{\pi_{11}\pi_{12}}{\pi_{12}\pi_{21}}$$

θ, 1'e eşit ise iki değişken arasında ilişki yoktur. θ, hem artış hem de azalış yönünde 1'den uzaklaştıkça ilişki kuvvetlenir. $1<\theta<\infty$ durumunda paydaki değişken için başarı odds'u, paydadaki değişkenden yüksektir. $0<\theta<1$ durumunda ise, paydaki değişkenin başarı odds'u paydadakinden düşüktür.

θ odds oranlarının örnekleme dağılımı, sağa çarpık bir asimetri verir. Bu durumdan kaynaklanan yorum güçlükleri, odds oranları yerine doğal logaritmaları(In θ ya da $\mathrm{Log}_e\,\theta$) yorumlanarak ortadan kaldırılır. θ=1 durumunda In θ=0, 0<θ<1 durumunda In θ<0, 1<θ<∞ durumunda In θ>0 olacaktır.

Üçüncü bir değişkene göre iki değişken arasındaki ilişki yapısı incelenmek istenirse, üçüncü değişkenin her bir kategorisi için diğer iki değişken arasındaki odds oranları hesaplanır. Bu odds oranlarına *ikinci dereceden(koşullu) odds oranı* adı verilir [2].

Üçüncü değişkenin her bir kategorisi için koşullu odds oranları birbirine eşit çıkarsa, üçüncü değişkenin diğer iki değişken üzerinde etkili olmadığı kabul edilir. Yani, üçlü etkileşim olmadığı sonucuna varılır. İkinci dereceden odds oranlarının hesaplanmasında, hangi değişkenin üçüncü değişken olarak alındığı sonucu değiştirmez. Bu özellik, değişkenler arasında bağımlı-bağımsız değişken ayrımının olmamasından ileri gelir.

Değişkenler arasındaki ilişki yapısının incelenmesinde, odds ve odds oranlarının yanı sıra kısmi odds ve kısmi odds oranlarının da önemli bir rolü vardır. Bir değişkenin tüm kategorileri için hesaplanan odds değerlerinin geometrik ortalaması, kısmi odds değerini verir. Bir başka ifade ile kısmi odds, koşullu oddsların ortalamasıdır.

Kısmi odds, değişkenlerden biri sabitken başarı durumunun başarısızlık durumuna göre şansını verir. Marjinal odds bir değişkenin tek başına odds değerini verirken, kısmi odds ikinci bir değişkenin sabitlenmesi koşulu altındaki odds değerini verir. Marjinal ve kısmi odds değerlerinin farklı çıkması durumuna, Simpson Paradoksu adı verilir. Hesaplanan marjinal ve kısmi odds oranlarının birbirleri ile karşılaştırılarak yorumlanması sonucunda, değişkenler arasındaki ilişki yapısı ortaya çıkarılır [4].

Benzer şekilde, üçüncü bir değişkenin tüm kategorileri için hesaplanan ikinci dereceden koşullu odds oranlarının geometrik ortalaması da, ikinci dereceden kısmi odds oranını verir. Kısmi odds oranı, üçüncü değişken sabitken, diğer iki değişken arasındaki ilişkiyi verecektir. Kısmi odds oranları da marjinallerinden farklı çıkabilir [2].

Ki-Kare Bağımsızlık Testleri

Değişkenler arasındaki ilişki yapısını incelemek amacıyla sıklıkla başvurulan testler, bağımsızlık testleridir. Bağımsızlık testleri, tesadüfi bir örneklemden elde edilen çapraz sınıflandırma tablosundaki n_{ij} göze frekansları ile tahmin edilen f_{ij} beklenen frekanslarının karşılaştırılmasına dayanır.

Bu testlerde H_0(sıfır hipotezi), n_{ij} göze frekanslarının, f_{ij} beklenen frekanslarına eşitliğini ifade eder. Bu hipotezi test etmede sıklıkla kullanılan test istatistikleri, Pearson χ^2 istatistiği ve Fisher'in G^2 benzerlik oran istatistiğidir [5].

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

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

2x2'lik çapraz sınıflandırma tablosu için yapılan bağımsızlık testlerinde, test istatistiğinin hesaplanabilmesi için beklenen frekanslar f_{ij}= n.π_{ij} = n.$\pi_{i+}\pi_{+j}$ işlemi ile bulunur.

Ki-kare dağılımlı tüm uyum iyiliği istatistikleri, büyük örneklem teorisine dayanır. Göze frekanslarının yeterince büyük olduğu hallerde her iki test istatistiği de aynı ki-kare dağılımına sahiptir ve yakın sayısal değerleri verirler. Dağılımın serbestlik derecesi (I-1)(J-1)'dir.

χ^2 ve G^2 istatistiklerinin ki-kare dağılımından uzaklaşmamaları için, beklenen göze frekanslarının en az 1 ve göze frekanslarının en fazla %20'sinin 5'den küçük olması istenir [5].

v_1 serbestlik dereceli bir ki-kare istatistiği ile, bundan bağımsız v_2 serbestlik dereceli bir başka ki-kare istatistiği söz konusu iken, bu iki istatistiğin toplamı v_1+v_2 serbestlik dereceli ki-kare dağılım gösterir. Bu özellik yardımıyla, serbestlik derecesi 1'den büyük olan ki-kare istatistikleri daha küçük serbestlik derecesine sahip bileşenlere ayrılabilirler. Böylece, bağımsızlık testlerinde, çapraz sınıflandırma tablosu parçalanarak belirli kategoriler arasındaki ilişki daha küçük boyutlu tablolarla incelenebilir.

χ^2 test istatistiği bölünebilme özelliğine sahip olmadığından bu amaçla kullanılamamaktadır. Fakat G^2, tam bölünebilme özelliğine sahiptir. Bu özellik, G^2 olabilirlik oran istatistiğine, Pearson χ^2 istatistiğine karşı üstünlük sağlar [1].

χ^2 ve G^2 test istatistikleri ile bu değerler için bulunan p kuyruk olasılıkları, H_0'ın reddi için yeterli kanıt bulunup bulunmadığını gösterirler. Fakat, gözelerin her biri için gözlenen ve beklenen frekansları teker teker karşılaştırmak, hipotez testi sonucunun hangi gözelerden kaynaklandığı ile ilgili ayrıntılı bilgiler verir. Bu amaçla *düzeltilmiş artık* adı verilen aşağıdaki değerler hesaplanır.

$$\frac{n_{ij} - f_{ij}}{\sqrt{f_{ij}(1-p_{i+})(1-p_{+j})}}$$

H_0 geçerli iken, düzeltilmiş artıklar standart normal dağılım gösterir. Artık değerlerinin mutlak değerce 2'den büyük olması, ilgili gözenin H_0'a uyumunun zayıf olduğunu gösterir [1].

III. LOGLİNEER MODELLER

Ki-kare analizi ile üç ya da daha fazla değişkenin iç içe çapraz tablolarının analizi yapılamamaktadır. Ancak ayrı ayrı IxJ tabloları biçiminde düzenlenerek analizler yapılabilmekte, ikili, üçlü, çoklu etkileşimler ve birlikte değişimler analiz edilememektedir. Loglineer analiz, hem ki-karenin uygulanabildiği Ix2 ya da 2xJ boyutlu tabloları, hem de ki-karenin yetersiz kaldığı üç ve daha büyük boyutlu tabloları modeller aracılığıyla analiz eden bir yöntemdir [6].

Loglineer modeller, Poisson ya da Multinomial dağılımlı veriler için geliştirilmiş lineer modellerdir. Bu modellerde kategorik değişkenin düzeyleri ile göze frekansları arasındaki ilişkiler araştırılır. Böylece bir kategorik değişkenler kümesi içinde ilişki ve etkileşim yapısı ortaya çıkarılabilir.

Fischer'in öne sürdüğü fikirlerden yola çıkarak, Birch (1963) tarafından ortaya atılan loglineer modeller, Goodman, Haberman ve Andersen'in katkılarıyla geliştirilmiştir. 1970'lerde sosyal bilimler alanında araştırma yapanlar tarafından sıkça kullanılmaya başlanan bu modeller, sadece ilişki yapısını çözmekle kalmayıp, aynı zamanda veriler için uyumlu bir modelin elde edilmesine olanak sağlaması yönüyle de, kategorik veri analizinde diğer teknikler arasında ön plana çıkmıştır [7].

Loglineer modellerin bir diğer özelliği, sınıflandırılan değişkenler arasında bağımlı-bağımsız değişken ayırımı yapmamasıdır. Loglineer analiz, değişkenler arasında nedensel ilişki olmaması temeline dayanır. Bununla birlikte nedensel ilişki analizleri için de, parametre yorumları farklı olmakla birlikte loglineer modelleri kullanmak mümkündür [8].

Tesadüfi bir örneklemden elde edilen gözlemlerin, iki kategorik değişken için IxJ boyutlu bir çapraz sınıflandırma tablosunda gösterildiğini düşünelim. X değişkeni satırda, Y değişkeni sütunda sınıflandırılmak üzere, X ve Y değişkenleri arasında bağımsızlık, $\pi_{ij} = \pi_{i+}\pi_{j+}$ eşitliği ile ifade edilir. Çarpım formundaki bu eşitliğin her iki tarafı için logaritma alınırsa, loglineer modellere adını veren toplam formundaki eşitlik ortaya çıkar. İki yönlü çapraz sınıflandırma tabloları için bağımsızlık durumunu ifade eden bu loglineer model şu şekilde ifade edilir.

$$In\pi_{ij} = \lambda + \lambda_i^X + \lambda_j^Y$$

Modelde λ genel etki terimi, i'inci satırdaki olasılığa ilişkin λ_i^X ana etki terimi ve j'inci sütundaki olasılığa ilişkin λ_j^Y ana etki terimi yer almaktadır. Modeldeki X ve Y, değişkenleri temsilen yazılan semboller olup, terimlerin üssü anlamında kullanılmamaktadır.

İki kategorik değişken için bağımsızlığı ifade eden H_0, bu loglineer modeli savunan hipotezdir. Modelin öngördüğü tahmin değerleri, $\hat{f}ij =$ n.$\hat{\pi}_{ij}$ değerleridir. H_0'ın geçerli olup olmadığı, Pearson χ^2 ya da G^2 benzerlik oran istatistikleri kullanılarak sınanabilir. Beklenen frekanslar, gözlenen frekanslara çok yakın çıkarsa, odds oranı 1 olur ve her iki test istatistiği de anlamsız çıkar.

İki yönlü bir çapraz sınıflandırma tablosu için, değişkenler arasında bağımlılık olduğunu gösteren loglineer model ise aşağıdaki gibi formüle edilir.

$$\text{Inf}_{ij} = \lambda + \lambda_i^X + \lambda_j^Y + \lambda_{ij}^{XY}$$

Yukarıdaki loglineer model ilk bakışta, varyans analizi modeline benzetilebilir. Fakat varyans analizi modelinde, bağımlı değişken bir gözlem skoru iken, loglineer modelde belirli bir sonucun gözlenme olasılığıdır. Bununla birlikte varyans analizinin tersine, loglineer analizde odak noktası ana etkileri gösteren tek değişken parametreleri değil, λ_{ij}^{XY} ikinci dereceden ilişki terimleridir [1]. Bu tür ilişki parametreleri, beklenen frekanslar hesaplanırken bir değişkenin diğer değişkenin kategorileri üzerinde etkili olduğunu gösterir [9-10].

Loglineer modellerde serbestlik derecesi, çapraz sınıflandırma tablosundaki göze sayısı ile modeldeki parametre sayısının farkı ile bulunur. İki değişken arasında ilişki olduğunu gösteren yukarıdaki loglineer modelde parametre sayısı, çapraz sınıflandırma tablosundaki göze sayısından fazla olduğu için serbestlik derecesi negatif çıkar, model tanımlanamaz. Modelin tanımlanabilirliğini sağlamak amacıyla, parametreler üzerine kısıtlar getiren bazı kodlama yöntemlerinin kullanılması gerekir.

Kullanılabilecek kodlama yöntemlerinden biri, parametrelerin etkilerini, her bir değişkenin belirli bir kategorisindeki etkilerden sapmalar şeklinde ifade etmektir. Regresyon analizinde de kullanılan bu kodlamaya *dummy kodlaması* denir. Genellikle ilk satır ve ilk sütuna ya da son satır ve son sütuna ilişkin parametre sıfıra eşitlenir. Parametrelerin etkileri bu gözeye göre yorumlanacağı için seçilen gözeye *referans gözesi* denir.

$$\lambda_1^X = \lambda_1^Y = \lambda_{11}^{XY} = 0 \text{ ya da}$$

$$\lambda_I^X = \lambda_J^Y = \lambda_{IJ}^{XY} = 0$$

i= 1, 2,...,I j=1, 2,...,J

Varyans analizinde de kullanılan bir diğer kodlama yöntemi ise, her bir etkiyi ortalama etkiden sapmalar şeklinde değerlendirerek, λ'ların toplamını sıfır kabul etmektir. Bu tür kodlamaya ise, *etki kodlaması* adı verilir.

$$\sum_i \lambda^X = \sum_j \lambda^Y = \sum_j \lambda_{ij}^{XY} = \sum_i \lambda_{ij}^{XY} = 0$$

i= 1, 2,...,I j=1, 2,...,J

Bu iki kodlama yöntemi ile aynı beklenen frekans değerlerine ulaşılmakla birlikte, hesaplanan parametre tahminleri farklı çıkmaktadır. Bu nedenle parametreler yorumlanırken, kullanılan kodlama yöntemi dikkate alınmalıdır [11].

İki yönlü çapraz sınıflandırma tabloları için, tüm ana etki ve etkileşim terimlerini içeren bu tür modellere, *doymuş(saturated) loglineer model* adı verilir [9]. Bağımsızlık modeli doymuş modelin $\lambda_{ij}^{XY}=0$ için özel halidir.

Modeldeki parametrelerin i ve j indisleri üzerinden toplamları sıfır olduğundan, bir sabit parametre, (I-1) sayıda λ_i^X ana etki parametresi, (J-1) sayıda λ_j^Y ana etki parametresi ve (I-1)(J-1) sayıda λ_{ij}^{XY} ilişki parametresinin hesaplanması yeterlidir. Bu durumda tahmin edilmesi gereken parametre sayısı, 1+(I-1)+(J-1)+(I-1)(J-1)=IJ'dir. Modelin parametre sayısı, çapraz tablodaki göze sayısına eşit çıktığından, verilere mükemmel uyum sağlar. IJ değeri mümkün olan maksimum parametre sayısını verir. Modele doymuş model denmesinin sebebi budur. Modelin uyum iyiliğinin sınanmasında, χ^2 ve G^2 test istatistikleri için serbestlik derecesi 0'dır.

Bir loglineer modelin ana etki parametreleri odds değerleri ile, ilişki parametreleri de odds oranları ile doğrudan ilgilidir. Her bir değişkenin iki kategorisi için λ_i^X ya da λ_j^Y ana etki parametrelerinin farkı, iki kategori arasındaki log-odds değerini verir. λ_{ij}^{XY} parametreleri de aşağıda gösterildiği gibi, log-odds oranlarının hesaplanmasında kullanılır [1].

$$In(\theta) = In\left(\frac{f_{11}f_{22}}{f_{12}f_{21}}\right)$$

$$= Inf_{11} + Inf_{22} - Inf_{12} - Inf_{22}$$

$$= \lambda_{11}^{XY} + \lambda_{22}^{XY} - \lambda_{12}^{XY} - \lambda_{21}^{XY}$$

λ_{ij}^{XY} parametreleri sıfıra eşit olunca, log-odds değeri de sıfır olur. Dolayısıyla odds oranı 1'e eşit çıkar. Bu sonuç, X ve Y arasında ilişki olmadığını gösterir.

Doymuş modeldeki parametrelerden bir ya da birkaçının hariç tutulduğu modellere *doymamış (unsaturated) loglineer model* denir. Doymamış modeller bazı parametreler üzerine konan kısıtlarla, daha sade ve kolay yorumlanabilir bir forma sahip olduklarından doymuş modellere tercih edilir [9]. Verilere en iyi uyum sağlayan doymamış model araştırılırken, modeldeki hiyerarşik yapıya dikkat edilmelidir.

Hiyerarşik yapı, yüksek dereceli bir terimin modelde bulunduğu hallerde, daha düşük dereceli tüm terimlerin de modelde yer aldığı türde ilişkileri ifade eder. Hiyerarşik bir loglineer modelde, bir λ parametresi sıfıra eşit ise, daha yüksek dereceli tüm λ parametreleri de sıfıra eşit olur. Aynı zamanda, bir λ parametresi sıfırdan farklı ise, daha düşük dereceli tüm parametreler de sıfırdan farklıdır [8].

Loglineer modellerin çözümünde kullanılan yöntemler ve paket programların hemen hepsi, modelin hiyerarşik olduğu varsayımına dayanır. Çünkü, bir etkileşim teriminin sıfıra eşit olması durumunda, daha yüksek dereceli etkileşim terimlerini yorumlamak anlamlı olmamaktadır.

Üç ve daha büyük boyutlu çapraz sınıflandırma tabloları için, değişkenler arasındaki "ilişki" terimi yerine yaygın olarak "etkileşim" terimi kullanılmaktadır. X, Y ve Z şeklindeki üç değişkenin mümkün tüm kategori kombinasyonları için, ixjxk boyutlu bir çapraz sınıflandırma tablosunda sınıflandırıldığını düşünelim. Böyle bir üç yönlü tablo için, doymuş loglineer model aşağıda verilmiştir.

$$\text{Inf}_{ijk}\, \lambda + \lambda_i^X + \lambda_j^Y + \lambda_k^Z + \lambda_{ij}^{XY} + \lambda_{ik}^{XZ} + \lambda_{jk}^{YZ} + \lambda_{ijk}^{XYZ}$$

En yüksek dereceli parametre dikkate alınarak, bu model $\{XYZ\}$ modeli biçiminde ifade edilir ve model parametrelerinin yorumu en yüksek dereceli terime (λ_{ijk}^{XYZ}) yönelik olarak yapılır.

Modeldeki tek değişken parametreleri değişkenlerin ana etkilerini verir. İkinci dereceden etkileşim terimleri, ikişerli olarak değişkenler arasındaki kısmi ilişkileri gösterir. Üçüncü değişkenin kategorilerinde diğer iki değişken arasındaki ortalama ilişki olarak yorumlanır. Modeldeki üçüncü dereceden etkileşim terimleri ise, üçüncü değişkenin kategorilerinde diğer iki değişken arasındaki ilişkinin farklı olabileceğini ifade eder. Bir başka ifade ile, kısmi ve koşullu ilişkiler arasındaki farkın büyüklüğünü ortaya koyar [9] .

Üç yönlü bir çapraz sınıflandırma tablosu için tahmin edilen model, üçüncü dereceden etkileşim parametresini içermezse, $\{XY, XZ, YZ\}$ şeklinde gösterilen bir doymamış model elde edilir. *Homojen ilişki modeli* adı verilen bu model, üçüncü değişkenin her bir kategorisi için herhangi iki değişken arasındaki koşullu odds oranlarının aynı olduğunu ifade eder.

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

İkili etkileşim parametrelerinin bir ya da bir kaçının modelden çıkarılması da mümkündür. Hiç bir etkileşim parametresi içermeyen, sadece ana etki parametrelerinden oluşan aşağıdaki model ise, değişken çiftleri arasında ilişki olmadığını gösterir. $\{X, Y, Z\}$ şeklinde gösterilen bu tür doymamış modellere *karşılıklı bağımsızlık modeli* adı verilir.

$$\text{Inf}_{ijk} = \lambda + \lambda_i^X + \lambda_j^Y + \lambda_k^Z$$

Loglineer modelin parametre tahminlerinin bulunmasında, en sık başvurulan yöntem, maksimum benzerlik(en çok olabilirlik) yöntemidir. Maksimum benzerlik yöntemi, ağırlıklı en küçük kareler, minimum ki-kare, minimum diskriminant bilgisi gibi alternatifleri arasında, olasılık dağılımı bilinen tesadüfi değişkenler için güçlü bir tahmin yöntemi olması nedeniyle tercih edilmektedir [9].

Farklı loglineer modellerin karşılaştırılması yoluyla, değişkenler arasında ilişkileri ifade eden parametrelerin anlamlılık testi yapılabilir. Karşılaştırılan iki modelden birinde test edilecek parametre/parametreler yer alırken, diğerinde yer almaz. Bu parametrelerin sıfıra eşitliğini ifade eden H_0'ın testi için, tam bölünebilme özelliğine sahip olan G^2 benzerlik oran istatistiğinden yararlanılır. İki modelin G^2 değerleri arasındaki fark, serbestlik derecesi, iki modelin serbestlik derecelerinin farkı olmak üzere test edilir. Kuyruk olasılığının küçük çıkması, H_0'ın reddedileceğine, yani incelenen ilişkinin anlamlı olduğuna işarettir. Kuyruk olasılığının büyük çıkması durumunda ise, iki modelin uyum iyiliği yönünden farkı olmadığı sonucuna varılır. Bu durumda, test edilen ilişki terimlerinin modele katkısı olmadığı görüldüğünden, bu terimleri içermeyen daha az parametreli model tercih edilir [12].

Çok yönlü tablolarda test edilecek model sayısı çok fazla olacağından, hipotezlerin sistematik bir yöntemle test edilmesi gerekmektedir. Bu nedenle, test edilecek hipotezleri anlamlı bir yorum çıkaracak şekilde gruplandırıp, sırayla test etmek en akılcı yoldur. Goodman bağımsızlık, koşullu bağımsızlık gibi hipotezlerin yorumlanabileceğini, diğerlerinin yorumlanabilir olma özelliği taşımadığını göstermiştir. Haberman ve Andersen tarafından geliştirilen bu sisteme göre, üç yönlü tablolar için test edilmesi gereken başlıca hipotezler aşağıda verilmiştir [7].

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

$$H_{XY} : \lambda_{ijk}^{XYZ} = 0 \;,\; \lambda_{ij}^{XY} = 0$$

$$H_{XZ} : \lambda_{ijk}^{XYZ} = 0 \;,\; \lambda_{ik}^{XZ} = 0$$

$$H_{YZ} : \lambda_{ijk}^{XYZ} = 0 \;,\; \lambda_{jk}^{YZ} = 0$$

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

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

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

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

Hipotezlerin testi, en yüksek dereceli parametreleri içeren ilk hipotezle başlar. Bir sonraki hipoteze geçilebilmesi için o aşamaya kadarki tüm hipotezlerin kabul edilmiş olması gerekir. Arka arkaya gelen iki hipotezden biri kabul edilirken, diğeri red ediliyorsa verilere en iyi uyan model, bu iki hipotezin arasında yer alır. Parametreler çeşitli biçimlerde kombine edilerek test edilir ve sonuçta verilere en uygun model oluşturulur.

Üç yönlü tablolar için ifade edilen loglineer modeller ve parametre yorumları, dört, beş,... yönlü tablolar için de genişletilebilir. Fakat değişken sayısı arttıkça, ilişki ve etkileşim terimlerini açıklamak karmaşık ve kolay yorumlanamayan bir hal almaktadır.

IV. UYGULAMA

Bu çalışmada 2002 Mayıs dönemine ilişkin LES sonuçları, üç yönlü bir çapraz sınıflandırma tablosu şeklinde düzenlenerek loglineer analizle değerlendirilmiştir. Aralarındaki ilişki yapısı incelenen üç kategorik değişken, *öğrenim durumu, LES puanı ve puan türü* olarak tanımlanmıştır. Öğrenim durumu değişkeni, LES'e giren kişilerin lisans, yüksek lisans ve doktora mezunu olma durumlarına göre üç kategoriye ayrılmaktadır. LES puanı değişkeni, 45 ve üstü ile 45'in altı şeklinde iki kategoriye ayrılmıştır. LES sonuçları sayısal puan, sözel puan ve eşit ağırlıklı puan şeklinde üç ayrı ağırlıklandırma ile hesaplandığından puan türü değişkeninin de üç kategorisi vardır.

2002 Mayıs döneminde LES sınavına giren kişilerin 339'u doktora mezunu, 16017'si yüksek lisans mezunu, 172054'ü lisans mezunudur. Bu kişiler arasında 45 ve daha fazla puan alan kişi sayısı 120745, 45'in altında puan alan kişi sayısı 67664'dür. Elde edilen üç yönlü çapraz sınıflandırma tablosu aşağıda verilmiştir.

		Puan Türü		
Öğrenim Durumu	***LES Puanı***	Sayısal	Sözel	Eşit Ağırlıklı
Doktora	≥45	82	89	85
	<45	31	24	28
Yüksek Lisans	≥45	3093	3375	3296
	<45	2246	1964	2043
Lisans	≥45	35547	37901	37277
	<45	21804	19450	20074

Verilerin analizinde SPSS for Windows 10.0 paket programı kullanılmıştır. Loglineer analizin ilk aşamasında ana etki, ikinci dereceden etkileşim ve üçüncü dereceden etkileşim terimlerinin anlamlılığı incelenmiştir.

Tests that K-way effects are zero

K	DF	L.R.Chisq	Prob	Pearson Chisq	Prob	Iteration
1	5	314638,889	,0000	326300,830	,0000	0
2	8	351,414	,0000	351,972	,0000	0
3	4	1,421	,8406	1,423	,8403	0

Yukarıdaki tablodan anlaşılacağı gibi üçüncü dereceden etkileşim terimlerinin istatistiksel olarak anlamsız olduğunu ifade eden sıfır hipotezi reddedilememiştir (0.8403). Bu durumda, loglineer modelin hiyerarşik yapısı gereği ikinci dereceden etkileşim terimlerinden en az biri anlamsız olacaktır. Dolayısıyla verilere en uygun loglineer modelin, en fazla ikinci dereceden etkileşim terimlerini içeren doymamış bir model olduğu anlaşılmaktadır.

Modelde yer alacak ikinci dereceden etkileşim parametrelerini tespit etmek üzere yapılan kısmi ilişki testi sonuçları aşağıda verilmiştir.

Tests of PARTIAL associations.

Effect Name	DF	Partial Chisq	Prob	Iter
ÖĞRENİM*LESPUANI	2	93.173	.0000	2
ÖĞRENİM*PUANTÜRÜ	4	.130	.9980	2
LESPUANI*PUANTÜRÜ	2	258.372	.0000	2
ÖĞRENİM	2	299479.868	.0000	2
LESPUANI	1	15159.046	.0000	2
PUANTÜRÜ	2	.000	1.0000	2

İkinci dereceden etkileşim parametreleri incelendiğinde, öğrenim durumu-LES puanı ve LES puanı-puan türü değişkenleri için kısmi ilişki parametreleri istatistiksel olarak anlamlı bulunmuştur. Fakat, öğrenim durumu-puan türü değişkenleri arasında anlamlı bir ilişkiye rastlanmamıştır.

Ana etki terimleri incelendiğinde ise, puan türü değişkeninin istatistiksel olarak anlamlı olmadığı görülmektedir.

Bu durumda, istatistiksel olarak anlamsız bulunan öğrenim durumu-puan türü etkileşim parametreleri, oluşturulacak doymamış loglineer modelde yer almayacaktır. Fakat kısmi ilişki testlerinde anlamlı bulunmayan puan türü ana etki parametresi, LES puanı-puan türü etkileşim parametrelerinin varlığı nedeniyle modele alınacaktır.

Aynı sonuç, doymuş loglineer modelin parametre tahminlerinden de görülebilmektedir.

Note: For saturated models.500 has been added all observed cells. This value may be changed by using the CRITERIA=DELTA subcommand.

Estimates for parameters.

ÖĞRENİM*LESPUANI*PUANTÜRÜ

Parameter	Coeff	Std. Err.	Z-Value	Lower 95 CI	Upper 95 CI
1	-.013828	.05832	-.23711	-.12814	.10048
2	.030466	.06085	.50067	-.0880	.14973
3	.000989	.03009	.03290	-.05799	.05997
4	-.010404	.03134	-.33201	-.07183	.05102

ÖĞRENİM*LESPUANI

Parameter	Coeff	Std. Err.	Z-Value	Lower 95 CI	Upper 95 CI
1	.199985	.04206	4.75482	.11755	.28242
2	-.136274	.02168	-6.28473	-.17877	-.09377

ÖĞRENİM*PUAN TÜRÜ

Parameter	Coeff	Std. Err.	Z-Value	Lower 95 CI	Upper 95 CI
1	.016953	.05832	.29069	-.09736	.13127
2	-.023501	.06085	-.38622	-.14277	.09576
3	-.009078	.03009	-.30170	-.06806	.04990
4	.011955	.03134	.38150	-.01947	.07338

LESPUANI*PUANTÜRÜ

Parameter	Coeff	Std. Err.	Z-Value	Lower 95 CI	Upper 95 CI
1	-.064263	.02936	-2.18854	-.12182	-.00671
2	.057818	.03062	1.88803	-.00220	.11784

ÖĞRENİM

Parameter	Coeff	Std. Err.	Z-Value	Lower 95 CI	Upper 95 CI
1	-3.43313	.04206	-81.6257	-3.51558	-3.3507
2	.538593	.02168	24.839	.49609	.58109

LESPUANI

Parameter	Coeff	Std. Err.	Z-Value	Lower 95 CI	Upper 95 CI
1	.359512	.02117	16.98052	.31802	.40101

PUANTÜRÜ

Parameter	Coeff	Std. Err.	Z-Value	Lower 95 CI	Upper 95 CI
1	.022086	.02936	.75216	-.03547	.07964
2	-.022395	.03062	-.73132	-.08242	.03763

Parametre tahminlerine ilişkin bu bilgiler, 0,05 anlamlılık düzeyinde değerlendirildiğinde, öğrenim durumu-LES puanı-puan türü üçlü etkileşim parametrelerinin anlamsız olduğu görülmektedir (Z-value: -0.23711, 0.50067, 0.03290, -0.33201). Aynı sonuç öğrenim durumu-puan türü etkileşim parametreleri (Z-value: 0.29069, -0.38622, -0.30170, 0.38150) ve puan türü ana etki parametreleri (Z-value: 0.75216, -0.73132) için de geçerlidir.

Geriye doğru eleme yöntemi sonucunda türetilen en iyi model, öğrenim durumu-LES puanı, LES puanı-puan türü, öğrenim durumu, LES puanı, puan türü parametrelerini içeren modeldir. Öğrenim durumu Ö, LES puanı L, puan türü P ile gösterilmek üzere elde edilen loglineer model aşağıda tanımlanmıştır.

$$\text{Inf}_{ijk} = \lambda + \lambda_i^{Ö} + \lambda_j^{L} + \lambda_k^{P} + \lambda_{ij}^{ÖL} + \lambda_{jk}^{LP}$$

Backward Elimination (p=.050) for DESIGN 1 with generating class

ÖĞRENİM*LESPUANI*PUANTÜRÜ

Likelihood ratio chi square = .0000 DF=0 P=1.000

If Deleted Simple Effect is	DF	L.R.Chisq Change	Prob	Iter
ÖĞRENİM*LESPUANI* PUANTÜRÜ	4	1.421	08406	2

Step 1

The best model has generating class

ÖĞRENİM*LESPUANI
ÖĞRENİM*PUANTÜRÜ
LESPUANI*PUANTÜRÜ

Likelihood ratio chi square=1.42085 DF = 4 P=0.841

If Deleted Simple Effect is	DF	L.R.Chisq Change	Prob	Iter
ÖĞRENİM*LESPUANI	2	93.173	.0000	2
ÖĞRENİM*PUANTÜRÜ	4	.130	.9980	2
LESPUANI*PUANTÜRÜ	2	258.372	.0000	2

Step 2

The best model has generating class

ÖĞRENİM*LESPUANI
LESPUANI*PUANTÜRÜ

Likelihood ratio chi square=1.55043 DF = 8 P=0.992

If Deleted Simple Effect is	DF	L.R.Chisq Change	Prob	Iter
ÖĞRENİM*LESPUANI	2	93.043	.0000	2
LESPUANI*PUANTÜRÜ	2	258.244	.0000	2

Step 3

The best model has generating class

ÖĞRENİM*LESPUANI
LESPUANI*PUANTÜRÜ

Likelihood ratio chi square=1.55043 DF = 8 P=0.992

SPSS paket programında loglineer modelin parametre tahminleri, değişkenlerin son kategorisini referans göze kabul eden dummy kodlamasına dayanmaktadır. Öğrenim durumu için *1:doktora mezunu, 2:yüksek lisans mezunu, 3:lisans mezunu*, LES puanı için *1:45 ve üstü, 2: 45'in altı,* puan türü için *1:sayısal, 2:sözel, 3: eşit ağırlıklı* olmak üzere, elde edilen loglineer modelin parametre tahminleri aşağıda verilmiştir.

	Parametre	Tahmin	Standart Hata	Z değeri
1	λ	9,9070	,0068	1449,98
2	$\lambda_1^{Ö}$	-6,6052	,1098	-60,14
3	$\lambda_2^{Ö}$	-2,2832	,0133	-171,99
4	$\lambda_3^{Ö}$	,0000	,	,
5	λ_1^{P}	,0838	,0093	9,00
6	λ_2^{P}	-,0324	,0096	-3,39
7	λ_3^{P}	,0000	,	,
8	λ_1^{L}	,6193	,0085	72,97
9	λ_2^{L}	,0000	,	,
10	$\lambda_{11}^{ÖL}$	,5355	,1264	4,24
11	$\lambda_{12}^{ÖL}$	,0000	,	,
12	$\lambda_{21}^{ÖL}$	-,1452	,0170	-8,56
13	$\lambda_{22}^{ÖL}$	,0000	,	,
14	$\lambda_{31}^{ÖL}$	,0000	,	,
15	$\lambda_{32}^{ÖL}$	,0000	,	,
16	λ_{11}^{LP}	-,1326	,0117	-11,32
17	λ_{12}^{LP}	,0497	,0119	4,19
18	λ_{13}^{LP}	,0000	,	,
19	λ_{21}^{LP}	,0000	,	,
20	λ_{22}^{LP}	,0000	,	,
21	λ_{23}^{LP}	,0000	,	,

Bu tahmin değerleri çapraz sınıflandırma tablosundaki her bir göze için, loglineer modelde yerine konarak aşağıdaki beklenen frekanslar hesaplanmıştır.

Table Information

Factor	Value	Observed Count(%)	Expected Count(%)
ÖĞRENİM	1		
LESPUANI	1		
PUANTÜRÜ	1	82(0.04)	82.10(0.04)
PUANTÜRÜ	2	89(0.05)	87.70(0.05)
PUANTÜRÜ	3	85(0.05)	86.20(0.05)
LESPUANI	2		
PUANTÜRÜ	1	31(0.02)	29.54(0.02)
PUANTÜRÜ	2	24(0.01)	26.30(0.01)
PUANTÜRÜ	3	28(0.01)	27.16(0.01)
ÖĞRENİM	2		
LESPUANI	1		
PUANTÜRÜ	1	3093(1.64)	3131.24(1.66)
PUANTÜRÜ	2	3375(1.79)	3344.97(1.78)
PUANTÜRÜ	3	3296(1.75)	3287.79(1.75)
LESPUANI	2		
PUANTÜRÜ	1	2246(1.19)	2225.39(1.18)
PUANTÜRÜ	2	1964(1.04)	1981.14(1.05)
PUANTÜRÜ	3	2043(1.08)	2046.48(1.09)
ÖĞRENİM	3		
LESPUANI	1		
PUANTÜRÜ	1	35547(18.87)	35508.08(18.85)
PUANTÜRÜ	2	37901(20.12)	37932.33(20.13)
PUANTÜRÜ	3	37277(19.79)	37284.00(19.79)
LESPUANI	2		
PUANTÜRÜ	1	21804(11.57)	21826.08(11.58)
PUANTÜRÜ	2	19450(10.32)	19430.56(10.31)
PUANTÜRÜ	3	20074(10.65)	20071.36(10.65)

- -

Table Information

Factor	Value	Resid.	Adj.Resid.	Dev. Resid.
ÖĞRENİM	1			
LESPUANI	1			
PUANTÜRÜ	1	-.10	-.01	-.0.1
PUANTÜRÜ	2	1.30	.17	.14
PUANTÜRÜ	3	-1.20	-.16	-.13
LESPUANI	2			
PUANTÜRÜ	1	1.46	.34	.27
PUANTÜRÜ	2	-2.30	-.54	-.45
PUANTÜRÜ	3	.84	.20	.16
ÖĞRENİM	2			
LESPUANI	1			
PUANTÜRÜ	1	-38.24	-.86	-.68
PUANTÜRÜ	2	30.03	.67	.52
PUANTÜRÜ	3	8.21	.18	.14
LESPUANI	2			
PUANTÜRÜ	1	20.61	.57	.44
PUANTÜRÜ	2	-17.14	-.49	-.39
PUANTÜRÜ	3	-3.48	-.10	-.08
ÖĞRENİM	3			
LESPUANI	1			
PUANTÜRÜ	1	38.34	.86	.20
PUANTÜRÜ	2	-31.33	-.69	-.16
PUANTÜRÜ	3	-7.00	-.15	-.04
LESPUANI	2			
PUANTÜRÜ	1	-22.08	-.61	-.15
PUANTÜRÜ	2	19.44	.55	.14
PUANTÜRÜ	3	2.64	.07	.02

Goodness-of-fit Statistics

	Chi-Square	DF	Sig.
Likelihood Ratio	1.5533	8	.9918
Pearson	1.5477	8	.9919

Görüldüğü gibi doymamış formda olan bu loglineer modelin, doymuş modelle karşılaştırıldığında anlamlılığı 0,992 olasılıkla red edilememiştir. Modelin verilere uyumu düzeltilmiş artıkların (adjusted residuals) mutlak değerce 1'i geçmemesinden de anlaşılmaktadır.

Modelin genel yapısının yanı sıra, parametre tahminleri kullanılarak kısmi odds ve kısmi odds oranları da hesaplanabilir.

$\lambda_1^{Ö}$ = -6,6052 değeri, öğrenim durumu değişkeninin 1'inci kategorisinin son kategoriye göre log-odds değerini verir. Bu değer, LES'e giren kişiler arasından tesadüfi olarak seçilen bir kişinin doktora mezunu olma şansının lisans mezunu olmaya göre exp(-6,6052) = 0,00135 olduğunu gösterir. Benzer şekilde, lisans mezunu olma durumuna göre, yüksek lisans mezunu olma şansı, exp($\lambda_2^{Ö}$)=exp(-2,2832) =0,10196'dır. Parametre değerlerinin negatif çıkması, lisans mezunu olmaya göre diğer iki kategorinin olasılığının düşük olduğuna işaret etmektedir.

LES puanı değişkeni için parametre tahminleri değerlendirildiğinde, rastgele seçilen bir öğrencinin LES'den 45 ve üstü puan alma şansı, 45'in altında alma durumuna göre exp(λ_1^{L})=exp(0,6193)= 1,8576'dır.

İkinci dereceden etkileşim parametreleri yardımıyla da log-odds oranları bulunabilir.

Doktora mezunlarının 45'in altında puan alma durumuna göre, 45 ve daha fazla puan alma şansı, yüksek lisans mezunları ile karşılaştırıldığında, exp($\lambda_{11}^{ÖL} + \lambda_{22}^{ÖL} - \lambda_{12}^{ÖL} - \lambda_{21}^{ÖL}$)=exp(0,5355+0,1452)=1.9 753, lisans mezunları ile karşılaştırıldığında exp($\lambda_{11}^{ÖL} + \lambda_{32}^{ÖL} - \lambda_{12}^{ÖL} - \lambda_{31}^{ÖL}$)=exp(0,5355)=1.7083'tür. Odds oranının 1'den farklı çıkması eğitim durumu ile alınan LES puanı arasında ilişki olduğunu göstermektedir. Ayrıca, 1.9753 değerinin, 1.7083'den büyük olması, doktora ve yüksek lisans mezunları arasındaki başarı farkının, doktora ve lisans mezunları arasındaki farktan fazla olduğu anlamına gelir. Yüksek lisans mezunları lisans mezunları ile karşılaştırıldığında exp($\lambda_{21}^{ÖL} + \lambda_{32}^{ÖL} - \lambda_{22}^{ÖL} - \lambda_{31}^{ÖL}$) =exp(-0,1452)=0,8648 çıkar. Bu odds oranının 1'den küçük çıkması, yüksek lisans mezunlarının LES'den 45 ve daha fazla puan alma şansının hem doktora, hem de lisans mezunlarından az olduğunu gösterir.

Sayısal ve sözel LES puanları karşılaştırıldığında, 45 ve daha fazla puan alma şansı exp($\lambda_{11}^{LP} + \lambda_{22}^{LP} - \lambda_{12}^{LP} - \lambda_{21}^{LP}$) = exp(-0,1326-0,0497) = 0,8333, sayısal ve eşit ağırlıklı puanlar karşılaştırıldığında exp($\lambda_{11}^{LP} + \lambda_{13}^{LP} + \lambda_{21}^{LP} + \lambda_{23}^{LP}$) = exp(-0,1326) = 0,8758, sözel ve eşit ağırlıklı puanlar karşılaştırıldığında ise, exp($\lambda_{12}^{LP} + \lambda_{13}^{LP} + \lambda_{22}^{LP} + \lambda_{23}^{LP}$) = exp(0,0497) = 1,0509'dur. Hesaplanan odds oranlarının 1'den farklı olması, puan türü ile LES puanı arasındaki ilişkinin göstergesidir. Bu sonuçlara göre, sayısal puan türünde 45 ve daha fazla puan alma şansının, sözel ve eşit ağırlıklı puan türüne göre düşük olduğu anlaşılmaktadır.

V. SONUÇ

LES sonuçları incelendiğinde öğrenim durumu, alınan LES puanı ve puan türü arasında üçlü etkileşim olmadığı tespit edilmiştir. Elde edilen loglineer modelin, doymuş modelle karşılaştırıldığında istatistiksel olarak anlamsız olduğunu ifade eden H_0, 0,992 olasılıkla reddedilememiştir.

Ana etki terimleri incelendiğinde puan türüne ilişkin parametreler istatistiksel olarak anlamsız bulunmuştur. Fakat puan türünden daha yüksek dereceli puan türü-LES puanı etkileşim teriminin anlamlı bulunması nedeniyle, loglineer modelin hiyerarşik yapısı puan türü ana etki teriminin de modele alınmasını gerektirmiştir.

Modeldeki ikinci dereceden etkileşim terimleri incelendiğinde, adayın öğrenim durumu ile LES puan türü arasında istatistiksel olarak anlamlı bir ilişkiye rastlanmamıştır. Bununla birlikte alınan LES puanı ile puan türü ve LES puanı ile öğrenim durumu arasında anlamlı ilişkiler olduğu saptanmıştır.

Puan türü ve alınan LES puanı arasındaki ilişki incelendiğinde, sözel puan türünde 45 ve daha fazla puan alma şansı, sayısal ve eşit ağırlıklı puan türlerine göre daha yüksek bulunmuştur. Başarı şansının en düşük olduğu puan türü ise, sayısal puandır. Bu sonuç, üniversiteye giriş sınavında da olduğu gibi sayısal bölümdeki soruların öğrencilere daha zor gelmesinden kaynaklanmaktadır.

Öğrenim durumu ile alınan LES puanı arasındaki ilişki incelendiğinde, doktora mezunlarının LES'den 45 ve daha fazla puan alma şansı, hem yüksek lisans hem de lisans mezunlarına göre yüksek bulunmuştur. Yüksek lisans mezunlarının lisans mezunlarına göre başarılı olma şansı ise daha düşüktür. Görüldüğü gibi, eğitim seviyesi ile LES'deki başarı düzeyi arasında doğrusal bir ilişki yoktur. Bu durum, LES'in bilgi birikimini değil, genel yeteneği ölçen bir sınav olmasından ileri gelmektedir.

YARARLANILAN KAYNAKLAR

[1] AGRESTI, A., **An Introduction to Categorical Data Analysis**, John Wiley & Sons., New York, 1996, ss.16-70, 145-173.

[2] HAGENAARS, J.A., **Loglinear Models with Latent Variables**, Sage Publications, Newbury Park, 1993, ss.3-13.

[3] DAVIS, J.A., *"Hierarchical Models for Significance Tests in Multivariate Contingency Tables: An Exegesis of Goodman's Recent Papers"*, **Sociological Methodology**, Vol.5., 1973-1974.

[4] SIMPSON, E.H., *"The Interpretation of Interaction in Contingency Tables"*, **Journal of the Royal Statistical Society**, Series B, Vol. 25, 1951.

[5] MOORE, D.S., *"Tests of Chi-squared Type"*, in D'AGOSTINO, R.B.; STEPHENS, M.A., **Goodness-Of-fit Techniques**, Marcel Dekker, New York, 1986, ss.64-65.

[6] ÖZDAMAR, K., **Paket Programlar ile İstatistiksel Veri Analizi**. Anadolu Üniversitesi Yayınları No: 1001, Fen Fakültesi Yayınları No: 11. A.Ü. Kütüphane ve Dökümantasyon Merkezi, Eskişehir, 1997, s.437.

[7] ANDERSEN, E.B., **Discrete Statistical Models with Social Applications**, North-Holland Publishing Company, Amsterdam, 1980, ss.162-191.

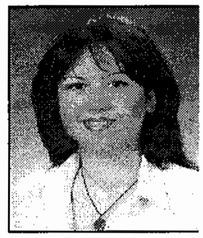
[8] LE, C. T., **Applied Categorical Data Analysis**, John Wiley & Sons., New York, 1998, ss.15-102.

[9] VERMUNT, J.K., **Log-Linear Models For Event Histories**, Sage Publications, California, 1997, ss.8-22.

[10] WILLEKENS, F., **Binary Data Analysis For Social Scientists**, Population Research Centre, Groningen, 1994, s.123.

[11] LONG, J.S., *"Estimable Functions in loglinear models"*, **Sociological Methods & Research**, Vol. 12, 1982.

[12] *"Tests of independence using multiway contingency tables in SPSS"*, (12 October 1999) [www document] Erişim: 31.07.2002.

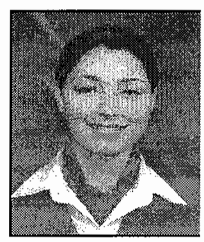


Dilek ALTAŞ

Marmara Üniversitesi, İ.İ.B.F.,
Ressam Namık İsmail Sk. No.1
Bahçelievler – İSTANBUL

dilekaltas@marmara.edu.tr

Dilek ALTAŞ has Ph.D. of Statistics at Marmara University Social Sciences Institute. She is Assistant Professor of Econometrics at Marmara University. Her research areas are bootstrapping methods, multivariate statistical analysis techniques.



Esen Zeren YILDIRIM

Marmara Üniversitesi, İ.İ.B.F.,
Ressam Namık İsmail Sk. No.1
Bahçelievler – İSTANBUL

eyildirim@marmara.edu.tr

Esen Zeren YILDIRIM is Research Assistant in Econometrics Department at Marmara University. Her research areas are multivariate statistical analysis techniques, applied statistics.